Universitat Autònoma
de Barcelona

HEALTH INEQUALITIES IN SPAIN:

Analysis of Structural Inequalities and the effects on Health and Mortality

Candidate:

Mathias Voigt

Supervisors:
Dr. Diego Ramiro Fariñas
Dr. Julio Pérez Diaz

Tutor:
Dr. Albert Esteve Pálos

Instituto de Economía, Geografía y Demografía (IEGD)
Centro de Ciencias Humanas y Sociales (CCHS)
Consejo Superior de Investigaciones Científicas (CSIC)

Department of Geography, Universitat Autònoma de Barcelona
Centre d'Estudis Demogràfics (CED)

2019

DESIGUALDADES DE SALUD EN ESPAÑA.

Análisis de las desigualdades estructurales y sus efectos sobre la salud y la mortalidad.

Candidate:

Mathias Voigt

Supervisors:
Dr. Diego Ramiro Fariñas
Dr. Julio Pérez Diaz

Tutor:
Dr. Albert Esteve Pálos

2019

Acknowledgements

Although a doctoral thesis appears to be a landmark in an individual researchers' career, often associated with long, lonely working hours and full of unavoidable suffering, the final product often bears the hallmarks of many people inside and outside of academia. I would not have had the courage to go out in the world and start such a journey if it was not for my parents, Thomas and Anke Voigt, who always supported me and who I could always rely and fall back on, when I needed a home, a piece of advice, or love and understanding. Although they might not always have understood what I was doing or why I was moving to Rome or Madrid, they never judged nor failed to show me that they trusted me. Danke, Mama! Danke, Papa!

I also owe a lot to my brothers, my grandparents, and my close friends who I left behind to go on this adventure. They always supported me, made it feel so easy to stay in touch, and always picked right up where we left off, the last time we saw each other. Thank you to all of you.

As ideas do not materialize themselves in a vacuum but are more likely a product of a long history of exchange with others, when writing this thesis, I benefited immensely from the ideas and nudges that I gathered from uncountable conversations and talks throughout my academic journey. As an aimless bachelor student, my first assistant job at the Rostocker Zentrum zur Erforschung des Demografischen Wandels got me interested in demography, but it was not until I worked for the Max-Planck Institute for Demographic Research when my dear co-workers, Jana Amtsberg, Maria Asmus, and Jonas Schöley, and I decided to apply for the Master in Demography.

Thanks to a great program and the excitement of my professor Roland Rau, I engaged further in the analysis of demographic problems and decided to take the next step and went to attend the European Doctoral School for Demography. Besides the great amount I learned about demography, the academic world, and myself during this time, it were my fellow EDSD-ler (or ACDC-ler) made these eleven months so special. They were not only an

immense source of support and knowledge but many of them became my friends. I owe special thanks to you, Andreas Höhn, for being an amazing friend, a source of trust and wisdom, and for great evenings with cheap Italian wine and smoked mozzarella.

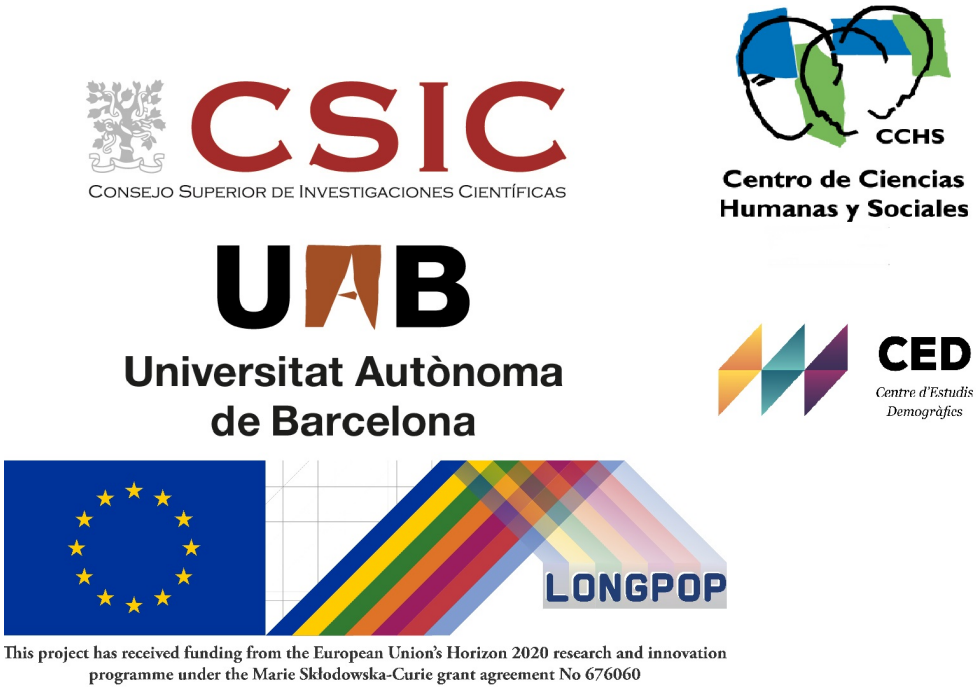
This dissertation would not have been possible without the continued support of my home institutions, the Instituto de Economía, Geografía y Demografía and the Centre d'Estudis Demográfics. I owe thanks to many of the staff and faculty at both institutions. First and foremost, I thank my directors Diego Ramiro Fariñas and Julio Pérez Diaz for the encouragement, patience, and support throughout the ups and downs of writing a doctoral thesis. Thanks to Antonio Abellán who generously shared his time to discuss my work. Thank you, Yolanda Casado and Pilar Aceituno for being such great sources of trust, for the free Spanish lessons, for helping me to make sense of so many cultural things, and for taking my mind off work when I needed it. Thanks to all the other great people, I have met here and spent long lunch hours with. Thanks to Yoan Molinero, Inmaculada Serrano, Alba Ayala, Tania Emerenko, Àngela Martinez-Carrasco, and Maria Dardoumpa for your commraderie and friendship. Through the LONGPOP project, I was lucky enough to be part of a network of so many bright, young researchers. Thank you to all my fellow ESRs!

I thank Albert Esteve Pálos, who gave me an institutional home to write my dissertation and was never too tired to answer my questions. Thank you, Socorro Sancho, for not losing your patience with me. Big thanks to Fran Viciana and the Institute of Statistics and Cartography of Andalusia, Spain, for providing generously their data and time. Thank you, Fran for answering my questions and for helping me to understand the rich data sources that were the basis for much of the work presented in this thesis.

Last but not least, thanks to the best companion for this journey, I can imagine. Thank you, Laura, for being my rock for the last four years. This thesis would have been impossible without your support, love, and belief in me.

This work was financially supported by the LONGPOP project which has received funding from the European Union's Horizon 2020 research and innovation program under the Marie Sklodowska-Curie grant agreement No. 676060 and the project LONGHEALTH CSO2015-69834-R funded by the Spanish Ministry of Economy and Competitiveness.

Disclaimer: This publication reflects only the author's view and the agency is not responsible for any use that may be made of the information it contains.



Abstract

Equality with regard to health and life spans is the basis for all other forms of equal opportunities. When structural inequalities determine the onset of chronic disability or the ages at death for different social group, they will naturally enhance themselves. In this thesis, these complex indirect association between one's social position, later-life health, and age at death are examined in the context of the elderly individuals in Spain and related to differences in age-specific mortality.

The three research articles, presented in this work, touch on different dimensions of inequality related to health and mortality. To quantify such effects, newly linked individual-level data is applied and the assessment of the long-term relationship between socioeconomic measures and mortality risks at different ages is conducted with different survival models.

Findings confirm persistent inequalities that especially affect the health and survival of the poorest and least educated. In addition to these repeatedly shown indirect associations between social position and health outcomes and the use of novel data sources, this thesis aims to expand classic analyses. The first article, aims to discuss the possible effects of recent changes in the formula to calculate retirement entry ages and pension benefits on inequalities in survival. The second article provides an alternative way to classify disability trajectories and reveals that socioeconomic differences have only a minimal effect on mortality after onset of disability. The third work contributes to the discussion of the effects of structural inequalities by addressing the importance of environmental risk factors. It proposes a tool to assess neighborhood "urbanicity" based on satellite images and finds that the social environment is more important for explaining differences in mortality than the degree of "urbanicity."

The complex and indirect relationship between health, mortality, and social inequalities is examined from different angles, the individual articles highlight how inequalities operate across the life course and how disadvantaged subpopulations can be identified and further be studied. As a whole, this thesis argues for continued exploration of linkable individual data sources and its application in social science and public health research, the use of more refined measures of disability and urban environment, and the incorporation of modern methodologies.

Desigualdades de salud en España.

La igualdad en salud y duración de la vida es la base de todas las demás formas de la igualdad de oportunidades. Cuando el inicio de una discapacidad crónica o la edad de la muerte resultan de las desigualdades estructurales entre diferentes grupos sociales, naturalmente esas desigualdades se retroalimentan. La compleja asociación indirecta entre la posición social, la salud en etapas posteriores de la vida y la edad al morir, se relacionan en esta tesis con las diferencias de mortalidad específica por edad, tomando como como población de estudio el conjunto de las personas mayores en España. Los efectos de los diferentes indicadores de desigualdad en salud y mortalidad se presentan en tres artículos de investigación. Se utilizan diversos modelos de análisis de supervivencia. El análisis estadístico de la información individual enlazada expresamente para este trabajo, está dirigido a medir la relación a largo plazo entre variables socioeconómicas y riesgos de mortalidad a diferentes edades.

A pesar de una elevada igualdad en términos de supervivencia, los hallazgos de los tres trabajos sugieren desventajas importantes en personas con menos recursos económicos y con menor nivel educativo. Además de estas repetidas asociaciones indirectas entre posición social y consecuencias para la salud, esta tesis tiene como objetivo ampliar el análisis clásico y contribuir al estudio de diferentes cuestiones de equidad social. El primer artículo, por ejemplo, analiza los posibles efectos de los cambios recientes en la fórmula para calcular la edad de entrada en jubilación y los beneficios de la pensión. El segundo proporciona una forma alternativa de clasificar diferentes trayectorias de discapacidad y además revela que las diferencias socioeconómicas tienen un efecto escaso sobre la mortalidad tras el inicio de la discapacidad. El tercer trabajo se suma a la discusión de los efectos de las desigualdades estructurales al abordar la importancia de los factores de riesgo ambiental, proponiendo una herramienta para evaluar la "urbanicidad" (o calidad urbana) de un área, basada en imágenes de satélite, y concluye que el entorno social es más importante que el propio carácter urbano de un área o barrio para el fenómeno analizado.

La investigación presentada en esta tesis contribuye de forma general a una mejor comprensión de cómo la salud, la mortalidad y las desigualdades sociales están conectadas. La relación compleja e indirecta entre estas tres dimensiones se examina teóricamente desde diferentes ángulos, mientras que los artículos concretos abordan aspectos específicos de esta asociación. Este trabajo destaca repetidamente cómo funcionan las desigualdades a lo largo del curso de vida y cómo se puede identificar y estudiar más a fondo las subpoblaciones desfavorecidas. En general, se aboga por mediciones más afinadas de la discapacidad y el entorno urbano que incorporen herramientas avanzadas de análisis estadístico. Reivindica la exploración de fuentes con datos individuales vinculados o enlazados y su aplicación para la investigación en ciencias sociales y salud pública.

Table of contents

List of figures

List of tables

Chapter 1

Introduction

My interest in the causes and consequences of structural social inequalities was probably sparked by the stories my grandparents and other family members told me about the social upheaval they lived through when Germany was reunified in 1990. These stories did not quite correspond with what public perception taught me. Growing up in the German Democratic Republic, in a political system that granted full-time employment to everyone, independent of their overall economic situation, my grandfather, one of my aunts, and two of my uncles were unprepared and hit hard when they lost their jobs during the restructuring of the German economy and the opening of the East to the West. While they were promised increased living standards and wealth, they instead faced hardship and uncertainty.

Back then, I did not quite understand how individual choices, financial constraints, and external events across the life course affect a person's health, or even the timing of his or her death. I had neither names nor concepts for what I saw until I came across them while a bachelor's student of Sociology at the University of Rostock. As I was learning about the sociological and economical macro theories, I became increasingly frustrated with the fact that these theories hardly ever provide a framework that allows explaining social phenomenon without a large number of exceptions. This lack of preciseness made them, in my opinion, weak tools and I longed for something that could be overthrown less easily. Naturally, I was drawn into by my Demography professor's passion for statistical solutions for problems on a population level during my last semester. Professor Roland Rau did not only sparked my interest in population science he also encouraged me to stick to demography after my Masters studies at the University of Rostock and apply for a position at the European Doctoral School of Demography (EDSD), a postgraduate program that would prepare me for most things to come. Throughout the program, I found myself constantly confronted with the pressure to find an identity as a researcher. If I wanted to continue on this path, I needed to decide what

I wanted to work on and how I wanted to invest my time and energy. Since we touched on almost everything within the field of population science, and various related topics, it was not until my final project that I came back to my interest in the impact and reproduction of structural inequalities. Returning to my initial interest also helped to be more confident and coherent on the search for a PhD position. In the spring of 2016, several of the EDSD teachers referred me to the ambitious Horizon 2020 project LONGPOP. From the beginning, I was intrigued by the position related to the individual project on health inequalities in Spain. Although I did not speak Spanish at the time and had little knowledge about the country's sociocultural background, I decided to apply and was accepted to the project several months later. An eventful journey began that culminated in the submission of this thesis about three years later.

As moving to a foreign country always comes with a certain period of familiarization, the first months in Madrid and at my new position were busy. Presented with the opportunity to use newly linked, longitudinal register data, provided by our partners at the Institute of Statistics and Cartography of Andalusia (IECA), it did not take long to familiarize myself with the Spanish material. I began to look forward to how my work could contribute to the analysis of inequalities in health and mortality. While working on my research proposal to apply for the PhD program in Demography at the Autonomous University of Barcelona, I formed the first ideas for how to approach the subject matter from different angles. Although I was proficient in many advanced demographic analysis techniques, I soon realized that I would need additional knowledge on methodologies and feedback from the research community. Because it was designed as international training network (ITN) and funded through the Horizon 2020 Program, my participation in the LONGPOP project offered the opportunity to further enhance my expertise. I attended as many as 9 training courses within the network and was provided funding to attend other workshops, conferences, and additional training outside the network.

Granted with this opportunity, I have been fortunate to present my work at the International Population Conference 2017 in Cape Town, the European Population Conference 2018 in Brussels and the 2018 and 2019 Annual Meetings of the Population Association of the Americas in Denver, Colorado and Austin, Texas. The feedback and ideas I gathered at these meetings in panel discussions and conversations with brilliant researchers were of great value for the overall development and completion of this work.

In the remaining part of this introduction, I will touch on the development and possible causes of structural inequalities in health and mortality in Spain and Andalusia. I will introduce the data used and, where necessary, methodology I employed in the succeeding chapters. While each chapter includes a detailed background section, the aim of this introduction is to provide a framework for the three research projects presented in this thesis. This includes an overview over the development of population and mortality indicators in Spain over the last 50 years. I further summarize previous research findings and sources of structural inequalities and their effect on average survival times and onset of disability. With focus on the population aged 50 and older, this background will include sections about the ideas that inequalities accumulate across the life course, how health and survival are measured and related at a population level, and why health inequalities persist in spite of collective efforts to reduce them.

Following the introduction, the research presented in the subsequent chapters will be structured in form of separated research essays resembling publishable articles which will be or have been submitted to peer-reviewed journals. When combined, they provide a comprehensive overview on health and mortality disparities within the elderly population of Spain.

All three essays were co-authored by other researchers whose ideas and valuable critiques kept me on the right path towards completion. I am very grateful for their collaboration, and it goes without saying that this work would not have been possible without their support. While I took the role of the lead author in all three projects, my supervisor Dr. Diego Ramiro and Dr. Francisco Viciana ensured that I had access to high quality data sources, including two linked versions of the Longitudinal Statistics on Survival and Longevity in Andalusia (BDLPA). The second essay on pathways through disability in the elderly population of Spain was made possible through Diego Ramiro´s effort to reach an agreement with the Spanish National Institute of Statistics (INE), to provide us with a linked data set between the National Survey on Disability, Autonomy, and Dependency (EDAD) and national mortality records. This essay further benefited greatly from the expertise of Prof. Antonio Abellán and Dr. Julio Pérez on the EDAD survey and the elderly population of Spain. Both provided me with invaluable knowledge on different classification schemes for disability and impairment as well as guidance for the work with this comprehensive source. The third essay on the effect of different exposures to environmental hazards in rural and urban areas of Andalusia benefited significantly from the support of my fellow LONGPOP colleague Dariya Ordanovich. She not only created and adjusted the maps that illustrate the results of this spatial analysis of mortality disparities, she also served as a wonderful source of information geographical measures and data visualization.

1.1 The Evolution of Longevity in Spain

1.1.1 The Rise of Life Expectancy in Spain

Decades of birth rates below replacement level and a historically unique improvement in mortality measures across all age groups have led to shrinking and heavily aging populations in most European countries over the last 50 years [140, 237]. While the population aging, induced by medical and social progress, can be translated as opportunity for millions of people to enjoy longer and healthier lives than the generations before them, the rise of average human life lengths also entails challenges through massive shifts in the age structure [56]. Though they trailed Northern and Western European countries in terms of longevity measures throughout most of the 20^{th} century, Spain became the country with highest life expectancy at birth in Europe in 2015 and is projected to overtake the top position in the world from Japan by the year 2040 [98]. Estimations based on data downloaded from the Human Mortality Database (HMD) indicate that the life expectancy of a Spanish newborn has increased on average by 12.4 years (12.48 years for girls and 12.32 years for boys respectively between 1964 and 2014 [134]).

A similarly strong increase in life expectancy at birth can be observed in all South European countries, depicted in the left panel of Figure 1.1. The black line, representing the Spanish trend, indicates the linear increase in life expectancy at birth began as early as 1950. There appears to be a catch-up period of even steeper growth following a period of stagnation in the 1940s, which is most certainly related to the disastrous state of the country after the Spanish Civil War [142, 234]. The development of life expectancy in the first period of the 20^{th} century is characterized by more fluctuations which can be traced back to periods of increased mortality during the civil war and pandemics like the Spanish Influenza of 1918/19. The end of the 19^{th} century marked a starting point for a period of rapid decrease in infant and child mortality in most European countries [30].

Though interrupted by wars and flu epidemics, medical and social progress led to the containment and elimination of various infectious and childhood diseases, the major cause of death at the beginning of the 20^{th} century. While prevention of infant and child deaths undoubtedly continued to be a major social and political goal, mortality improvements in older age groups also increasingly contributed to continuation of the growth in the average length of life [192, 307]. The steep increase in life expectancy at age 65 from the beginning of the 1970s, shown in the right panel of Figure 1.1, underlines that this shift also appeared in Spain.

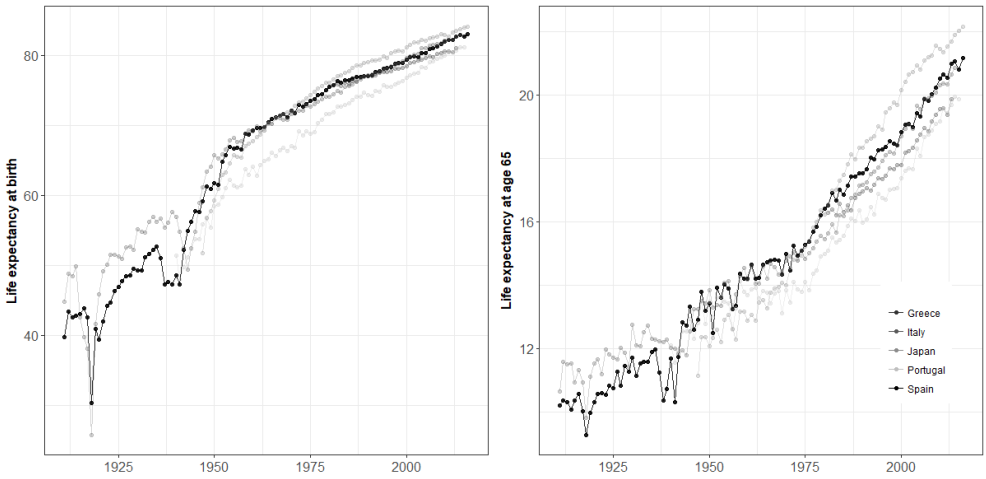


Fig. 1.1 Life Expectancy at Birth (left) and at age 65 (right), Spain (highlighted) and other Southern European countries between 1915 and 2016 (Calculations based on data downloaded from the Human Mortality Database)

Although the decrease in old age mortality has accelerated since the 1970s, the rate of increase in life expectancy at birth appears to have slowed down in the second half of the century. This does not reflect a slowed reduction in age specific death rates but merely a changed character of the process. Increased efficiency of medical treatments and behavioral changes, most importantly smoking cessation, led to an accelerated decrease in mortality rates due to non-communicable diseases, particularly affecting the population 60 and older. Since the 1980s, the average rate of increase in life expectancy at age 60 was more than double that of the period between 1950-80.

1.1.2 Lifespan Variability and Compression of Mortality

These last three decades of decreasing annual death rates in all age groups have led to compression of ages at death. Today, the majority of deaths in most high and middle-income countries occurs in a small age range between 85 and 95 [192]. To characterize these mortality developments in a different way, one can compare the age distribution of deaths and period survival curves over time. Using the yearly death counts and population exposure provided by the HMD, one can calculate the proportion of deaths at age x in year y, denoted as d_x, and the probability of survival for someone aged x (life table ${}_tp_x$). Before calculating the life table measures, I smoothed the yearly single age death counts with P-Splines [46]. Although Spanish death count data is generally reliable and of good quality, in two specific periods, during the Spanish civil war and immediately after, reporting on timing of deaths

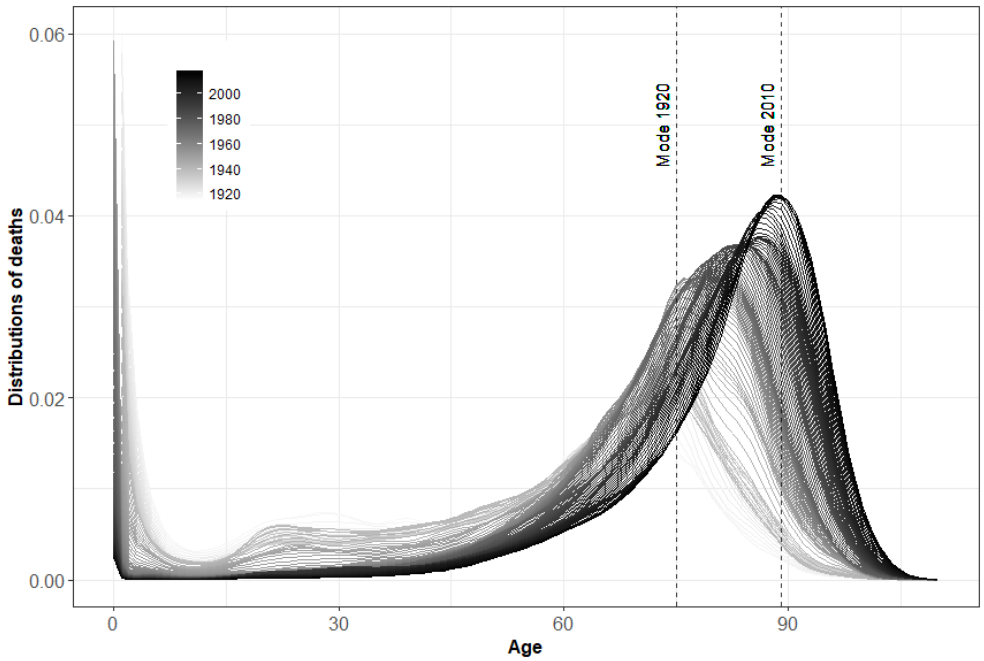


Fig. 1.2 Age-specific Distribution of Deaths over Time, Spain

appear to be delayed according to the *Movimiento Natural de la Poblócion* (MNP), the annual series of Spanish Vital Statistics. Prior to 1960, there were also problems due to age heaping, which affected females more than males [77].

The HMD recommends adjusting the data pre-1975 to prevent this year-to-year fluctuation and age heaping from affecting such calculations [115]. I used the R package *MortalitySmooth* which applies p-splines to smooth the death counts and exposures before creating period life tables for the 101-year period between 1915 and 2016. Extracted smoothed single age probabilities of dying are displayed via the density function in Figure 1.2. Darker shades represent more recent years.

In the early years, there the aforementioned reduction of infant and child mortality is visible. In the middle of the century, the modal age at death (not accounting for infant mortality) was then postponed into higher ages. The two vertical dotted lines represent the shift in the modal ages at death. The age with the highest estimated number of deaths increased by 14 years from age 75 in 1920 to age 89 in 2010. The latest period is characterized by a compression of deaths around the mode, a development associated with approaching a biological maximum age [216, 83]. If such a maximum exists, we have not reached it yet nor do predictions indicate that we will reach it soon [173, 210]. Another way to look at

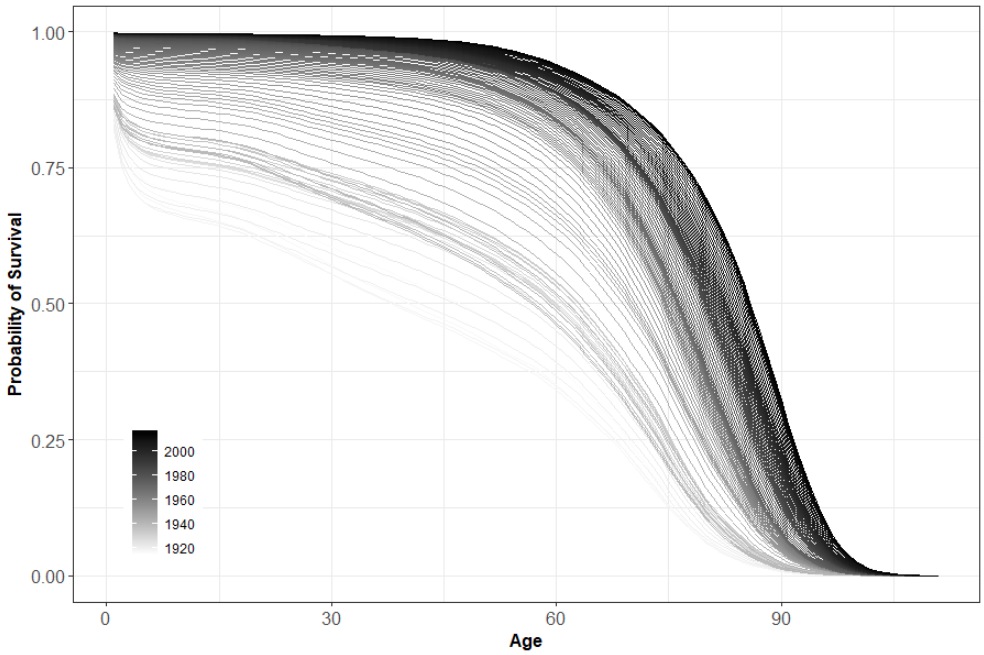


Fig. 1.3 Survival Probability by Age and Year for the Total Population of Spain (Calculations based on data downloaded from Human Mortality Database)

compression of deaths is the rectangularization of the survival curve over time, depicted in Figure 1.3. The apparent formation of a right angle at the top right corner of the graph indicates not only an increase in survival probabilities over time but growing statistical certainty about the timing of deaths [e.g. 184]. In other words, someone born today can expect to live 80 to 85 years with a high degree of certainty if the conditions (including the population structure) remain constant. There are numerous measures to quantify uncertainty, or in other words, the variability surrounding average life spans. Most of these measures can be directly derived from the life table [50, 296, 280, 151]. Figure 1.4 represents the trajectory of a lifespan variability measure, "e-dagger" ($e^{\dagger}$), which is calculated based on smoothed male and female life tables from the HMD. $e^{\dagger}$ can be interpreted as life expectancy lost due to death, similar to the concept of Potential Life Years Lost [108]. Mathematically it can be expressed as in equation 1.1.

$$e^{\dagger} = \int_0^{\omega} e(x,t) f(x,t) dx \tag{1.1}$$

Where $e(a,t) = \int_a^{\omega} \frac{l(x,t)dx}{l(a,t)}$ is remaining life expectancy at age a and time t and $l(a,t) = \exp(-\int_0^a \mu(x,t)dx)$ gives the probability of survival to age a and $\mu(a,t)$ denotes the age-specific hazard of death [296]. The standard deviation above the late modal age at death

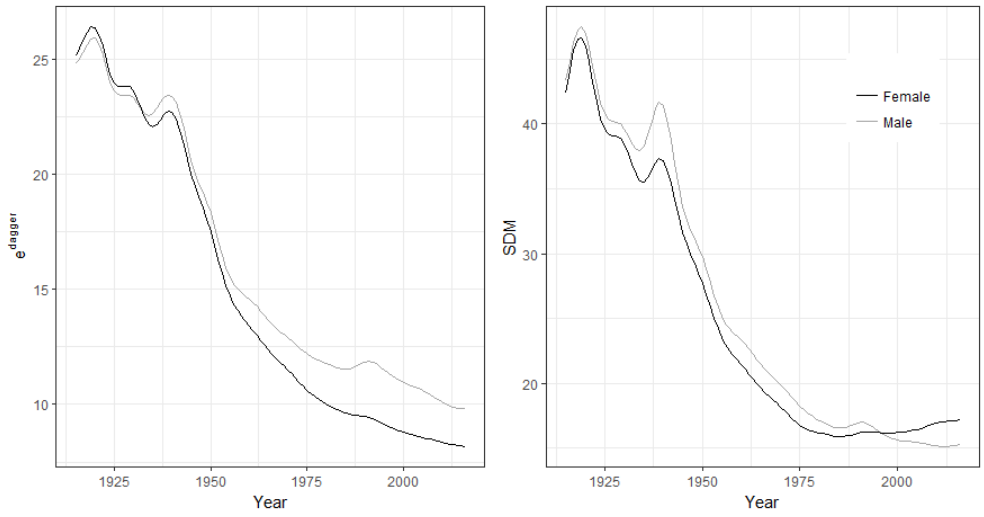


Fig. 1.4 Lifespan disparity measures by sex over time, Spain; left: $e^{\dagger}$; right: Standard deviation around the modal age at death (Calculations based on data downloaded from Human Mortality Database)

is another helpful measure of dispersion of life lengths. Although it does not differentiate between deaths occurring prematurely or late in life, it helps to quantify the concentration of deaths around the mode [50, 151]. The trajectories of both measures suggest that lifespan variability in Spain evolved very similarly for males and females before 1960. From then on, it appears that the decrease in female disparity accelerates while the male curve plateaus between the late 1970s and mid-1990s. The second variability measures, the standard deviation around the mode (SDM), accounts for the concentration of deaths around the modal age at death. It can be expressed as in equation 1.2.

$$SDM(t) = \sqrt{\int_0^{\omega} [a - M(t)^2] d(a,t) da}, \tag{1.2}$$

Where M refers to the mode at year y and ω to the highest age attained in the population [50]. Parallel to the standard deviation around the mean, both measures are expressed in units of years [308]. In contrast to $e^{\dagger}$ the variability measured through the SDM does show a divergence of male and female values but a cross over in the late 1990s, possibly related to the uptake in female smokers [262]. Lifespan variability indicators are a mostly descriptive measures of the shape of the distribution of lifespans [59]. Similar to life expectancy, these measures refer to population averages and are not designed to capture inequalities within populations. Differences in the shape of the distributions of deaths and average lifespans, however, can be partly interpreted as a manifestation of all social processes that directly and indirectly affect health and mortality. A growing body of research comparing life expectancy

and variability at the subnational level suggests the existence of persistent inequalities are due to differences in education, occupational class, and other wealth indicators. These analyses suggest that groups with lower levels of education and lower average income are more likely to face substantially lower average life expectancy and more uncertainty surrounding their age at death [215, 292, 55].

While the importance of reducing educational differences in average life lengths is self-evident, it appears to be less obvious to address differences in life span variability. However, large between-group inequalities regarding uncertainties of average life lengths can affect future mortality developments in two ways. From a population perspective, larger heterogeneity in the ages at death implies greater group differences in population health. These health differences are likely to be enhanced through larger investments by higher educated and wealthier groups who are assumed to claim and use additional resources to facilitate their healthy life styles and, thus, further reduce their variability in ages at death relative to other subpopulations [259, 40]. At an individual level, we tend to project our own survival based on experienced mortality within our family and friend groups [135]. These subjective expectations then become the baseline for important decisions affecting different life domains, including commitment to healthy lifestyles, economic investments, and retirement decisions [136]. Aggregated at a subnational level, those groups experience a higher number of premature deaths and, thus, face larger uncertainty. Consequentially, they will not invest in many life domains to the same extent as groups with low levels of uncertainty. The result is a self-reinforcing process resulting in further increasing disparities [259].

In one of the few analyses of this kind in Spain, Permanyer et al. link the Spanish Encuesta Sociodemográfica with newly released mortality files by the Spanish Statistical Office (INE) to study educational mortality differences over the last 50 years. They observe that, though all educational groups gained years of life expectancy, the trends diverged over time. Higher educated groups gained extra life years at a much higher rate than their lower educated counterparts did. Similar to the aforementioned findings in other countries, uncertainty was higher in lower educated groups [227].

1.1.3 The Challenges of Population Aging

Decades of accelerated mortality reduction and low birth rates have naturally led to changes in the population structure in most European countries. The resulting population aging appears to challenge the very social security systems that facilitated the unprecedented mortality improvements in the first place [56]. The growing share of the elderly who are potential

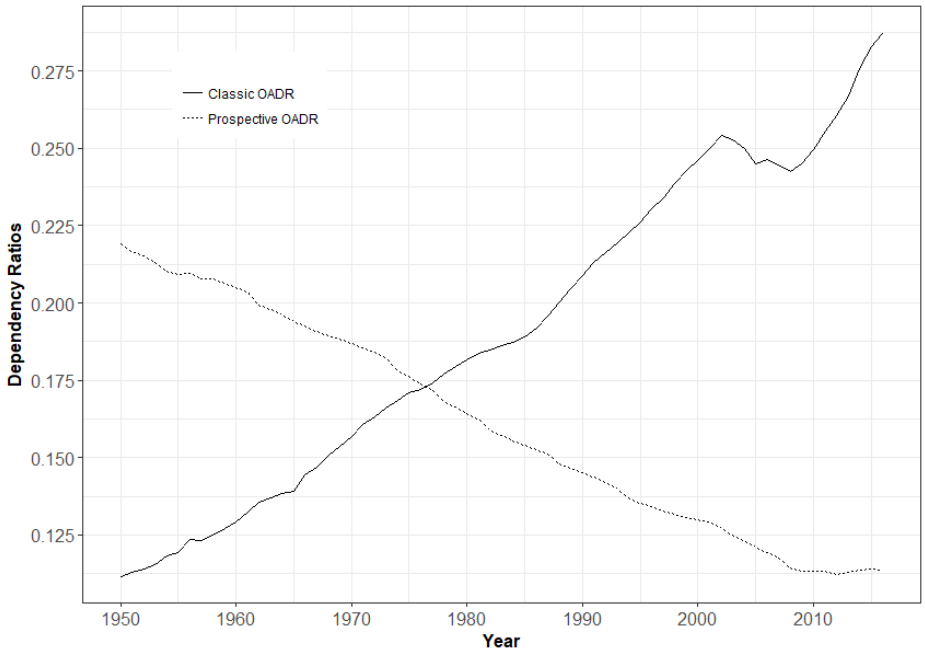


Fig. 1.5 Old-Age-Dependency Ratio and Prospective Dependency Ratio, Spain (1950-2016)

receivers of welfare benefits increases pressure on the shrinking working age population to contribute to social security systems. The National Statistical Office of Spain (INE) has estimated that the country will experience population aging more rapidly than most North or Central European countries [141]. The Old-Age-Dependency Ratio (OADR) is a measure that is often applied to quantify population aging over time. It can be expressed as in equation 1.3

$$D = \frac{N_{65+}}{N_{15-64}} \tag{1.3}$$

,where N_{65+} is the theoretically receiving population in year y, while N_{15-64} represents the subpopulation contributing to social security in that same year. Although this ratio has been criticized as inaccurate, it serves well as a descriptive overview [203]. Calculated based on age-specific population exposures extracted from the HMD, the development of the OADR over the last 45 years is presented in Figure 1.5. The graph suggests that the OADR has constantly increased with the exception of a short period of fluctuation in the early 2000s, when Spain experienced record immigration during an economic boom period.

The INE has projected that the ratio will further increase until it reaches a peak value of 0.5 in 2050. In other words, the number of potential contributors per receiver of old age pension benefits is expected to decline from three to two within 30 years. Due to demographic momentum, neither an increase of migration nor an abrupt increase in birth rates can prevent

population aging in Spain [89]. Consequently, there is scientific and economic interest in the extension of artificial age limits, which, for example, determine at what age a person is eligible for retirement benefits. To justify the extension of working lives, it is necessary to account for the quality of life within the extra years of life expectancy gained over the last several decades. When those extra years are generally spent in good health, it is easier to argue for a higher legal age to receive retirement benefits.

Indeed, we find that 65-year-olds today generally have higher remaining life expectancy and are healthier than their counterparts are in earlier generations. Traditional aging measures like the ODR do not capture these changes in population health. Although still widely used, these classic dependency ratios should be combined with forward-looking aging measures, like the idea of prospective aging. Instead of using chronological ages as markers for old age, Sanderson and Sherbov introduced this measure to estimate how much a population has aged over time based on relative changes in remaining (disability-free) life expectancy with reference to an index year [257]. Different from remaining life expectancy, proposed by Norman Ryder, prospective aging is insensitive to changes between period and cohort mortality. In period life tables, it can be calculated as shown in Formula 1.4

$$e_p(A_p, S) = e_p(a, I) \tag{1.4}$$

, where $e_p(A_p, S)$ refers to the prospective age of a person in the standard year S and $e_p(a, I)$ refers to the prospective age of a person of retrospective age a in the index year I [see 256]. I calculate the prospective OADR for Spain between 1950 and 2016 based on remaining life expectancy at 65 (e_{65}) in 1950, the index year. The trajectory over the last 65 years is depicted by the dotted line in Figure 1.5. While the classic OADR uses a fixed age at the onset of "old age dependency," here the legal eligibility age for receiving full retirement, the prospective OADR is based on the age at which the average person can expect to live as long as someone who reached the retirement entry age in the index year 1950. According to the previously mentioned smoothed life tables for the Spanish population, the remaining life expectancy at age 65 in 1950 was 13.18 years. Thus, the prospective age at the onset of "old age dependency" is the age in a given year at which the remaining life expectancy was 13.18 years. In contrast to the classic OADR, the prospective age measures capture the remarkable improvements in life expectancy and indicate that the dependent share of the population has decreased. The trend of the prospective dependency ratio has only begun to plateau in the last 5 years. Given the large share of the population aged 65 living independent and active lives, prospective age based indicators may give us a more realistic view on population aging

[256]. While most alternative aging measures provide a more positive look into the future, when based on average life tables, they tend to neglect heterogeneity within a population.

1.2 Longevity and Health

1.2.1 How to relate morbidity patterns to population dynamics?

A theoretical link between population dynamics and patterns of disease and disabilities was posited by Abdel Omran in 1971 [217]. In his "Theory of Epidemiological Transition," he proposed that changes in the causes of death distribution and related disease patterns are the main forces that drive population aging and other changes in various social contexts. He argues that 20th century gains in life expectancy were made possible by the eradication and containment of deadly infectious diseases. According to his essay, we had reached the "The Age of Degenerative and Man-Made Diseases [217]" and that life expectancy would increase until it reaches age 50. What he could not know at the time is how medical and social progress would affect future diseases and cause of death patterns. While he observed that heart and cardiovascular diseases have become the major cause of death in the Western world, nobody could have predicted the vast reductions in cardiovascular mortality and the continued rise in life expectancy over the last decades. To account for these reductions since the 1960s ("cardiovascular revolution"), Vallin and Meslé proposed a third stage to Omrans original framework which highlights the increasing number of deaths due to neurodegenerative diseases and cancer [286, 287]. While high-income countries are suggested to have reached a succeeding transition stage, the lack of time series data for causes of death data used to make it difficult to test Omrans assumption for most of the rest of the world. Recent analysis of global cause of death distributions indicate that mortality patterns differ strongly between wealthier and poorer regions of the world. In contrary to the assumed convergence, researchers find that regional differences in life expectancy tend to become larger [5].

1.2.2 Are gained life years healthy years?

As populations age at ever-faster rates, questions about the quality of the gained years in life expectancy have become increasingly important. Will we spend these extra years in generally good health, or are we facing extended periods of morbidity? Are mortality and morbidity rates declining at similar rates? Can we assure quality care to a growing group of elderly?
In 1980, James Fries proposed the concept of morbidity compression, which he related to the postponement of chronic disease and disability onset closer to the age of death for a

person. Fries stated that chronic illnesses, which account for more than 80% of all deaths, are "inescapably linked with eventual mortality [102]." These chronic illnesses and disabilities are thus identified as markers for possible compression or expansion of morbidity [102]. His framework and the adaptations to it, like dynamic equilibrium [182], laid the basis for assessments on whether additional life years are spent morbidity-free or not. Although criticized for its sole focus on chronic conditions [24], Fries's framework gives a starting point for the evaluation of a complex reality.

When describing morbidity trends over time, we combine various individual life course trajectories that range from individuals with fatal heart attacks at age 50, to those who live up to 100 and die asymptotically from multiple minor diseases or defects [112]. To test the compression of morbidity hypothesis, it is also often assumed that once a disability has occurred, there will be a steady and linear decline in health. Often due to the lack of longitudinal information and the convention to use the onset of first disability, the possibility of multi-phasic trajectories is hard to analyze. In other words, we do not always allow for successful recoveries or relapses from treatments [103]. Even if one accepts the assumption of linear decline of population health, there are multiple scenarios leading to compression or expansion of morbidity. In the original approach, it was assumed that mortality would asymptotically approach a maximum age while the average onset of morbidity would be delayed to higher ages. As these two hypothetical onset ages can increase or decrease at different rates, we are still left with a number of theoretical scenarios [see 24]. Starting from the baseline scenario or the current onset age of morbidity and age at death, it is assumed here that mortality improvements will continue to postpone the average age at death independent of the scenario of morbidity delay. The uncertainties about future scenarios regarding the onset of morbidity are rooted in the difficulties to define health and measure it at a population level. While death is a clearly defined concept and information on deaths are routinely collected at statistical offices, it is far more complicated to assess the health of a population or subpopulation. Even if health is solely defined as absence of disease or chronic condition, when diagnosed, many conditions can range from mild to incapacitating. Thus, one would need to select a threshold value to distinguish a healthy from a sick person. Such an exercise becomes increasingly difficult the larger the variety of considered conditions [35].

Instead of counting diagnoses, the use of impairments of functioning at activities of daily living (ADL) and instrumental activities of daily living (IADL) have been established as a proxy for morbidity at older ages [138, 155, 269]. Most commonly, ADLs include bathing, dressing, eating, transferring, and toileting while IADL refer to periodic routines often

including meal preparation, shopping, housekeeping, laundering, using the phone, managing medications, managing money, and using transportation [167]. Once it is decided how to define disability or health, one can use age-specific onset rates to estimate disability-free or healthy life expectancy (DFLE). The single most important method to estimate life expectancy free of disability was developed by Sullivan almost 50 years ago (1971) [277]. Although a formal justification of his index was not provided until the early 2000s, his method for combining mortality and morbidity rates has been the most widely used approach among applied researchers and governments agents for estimating DFLE [138]. The estimation of the index requires age-specific cross-sectional disability rates and a period life table, ideally for the same time period. Sullivan suggested to split the life table L_x, the person years lived in a given age interval, into the proportion with and without disability. One does so by multiplying the original Lx values with the sample fraction of disabled respondents from a survey. From there on, one would use the "new" Lx_i value to calculate the Tx_i and finally the ex_i, which represents the life expectancy at age x without disability. Expressed mathematically the Sullivan index could look like follows.

$$e_x^i = \frac{1}{l_x} \sum_{i=x} (1 - \pi_i)_{n_i} L_i, \tag{1.5}$$

where π_{n_i} is the fraction of disabled survey respondents. π_{n_i} will vary depending on the sampling method of the survey [138]. To estimate DFLE for Spain, I first calculate age-specific disability rates from the *Encuesta Sobre Discapacidades, Autonomía Personal y Situaciones de Dependencia* (EDAD), a nationally representative survey study focused on disability and autonomous living in Spain. The study design and population characteristics will be introduced in a later chapter. The EDAD data allows me to extract age-specific fraction of disabled and dependent individuals. I define an individual as disabled when she has stated difficulties in one or more of 13 pre-selected IADL and ADL. The list of included activities can be found in Figure (later). Being in dependency refers to individuals who stated that they need help the of a personal caretaker to perform one or more of the 13 activities.

I use the HMD life table from the survey year 2008 to estimate DFLE and the dependency-free life expectancy (DepFLE). Table 1.1 below shows the DFLE and DepFLE next to the overall life expectancy for selected ages in 2008. Given the different average health trajectories for men and women, results are stratified by sex [212]. Although it is just a snapshot, estimates in Table 1.1 indicate the existence of a health –survival paradox between the sexes. While women have higher life expectancy at all ages, their life expectancy free of disability appears to decrease at a lower rate compared to men. Between the ages 65 and 75, the two trend lines

cross and after age 75, female DFLE is lower than male DFLE. The same age pattern occur for the life expectancy free of dependency, indicating that women when compared to men, spend more relative time in disability or receiving care. Although women appear to be less healthy at higher ages, their average life spans are substantially longer compared to their male counterparts. Such a health-survival paradox was found in other vastly aging populations and piques interest as there are various contributing time varying factors including differences in the uptake of tobacco smoking [212, 291].

Table 1.1 Life expectancy, DFLE, DepFLE for both sexes and selected ages, Spain (2008)

	Females			Males		
Age	LE	DFLE	DepFLE	LE	DFLE	DepFLE
50	35.576	27.693	29.034	30.161	26.579	27.481
65	21.846	14.860	15.878	17.811	14.613	15.322
75	13.381	7.449	8.159	10.749	7.828	8.393
80	9.723	4.550	5.084	7.843	5.121	5.596
90	4.528	1.152	1.424	3.832	1.559	1.872

1.3 How inequalities in health and mortality arise and why do they persist?

There is strong evidence suggesting that health and age specific mortality are closely associated with income, wealth, education level, gender, ethnicity, and occupation. Numerous analyses confirm the existence of health and mortality advantages for better educated and wealthier individuals over their less wealthy and low educated counterparts [176, 183, 200, 215, 180, 177]. These disparities or health inequalities can be understood as presumably avoidable group differences, which are caused by structurally restricted access to key determinants of a healthy life style including quality education and fair pay [36]. The association between socioeconomic status (SES) measures and health or mortality outcomes can be found in societies around the world regardless of the extent of income inequalities or differences in the access to health care. Although modern welfare fare systems in most European countries were designed to reduce socioeconomic inequalities and its consequences, inequalities in mortality have been reported to increase in many countries. Their situation is even more paradoxical as there appears to be no association between the intensity and amount of welfare policies and the magnitude in inequalities with respect to health [175]. Such an evolution of the health gap, despite efforts to reduce poverty, exploitation, and social inequalities piques the interest of researchers in the mechanisms behind the development of health and survival inequalities [187].

As health inequalities ultimately derive from structural social differences, most explanations are based on an understanding of social stratification and its components, which include social mobility, the distribution of resources to social strata, and the changing value of resources [175, 221]. Among the approaches to explain increasing inequalities in health and mortality are the life course perspective, increased homogeneity of social groups in terms of health behavior, and the increased importance of consumption behavior due to shifts in the epidemiological regime.

In a simplified framework, the socioeconomic position and other societal conditions are understood to shape the more "downstream" determinants such as health-related behaviors or the exposure to environmental risk factors, which are more directly associated with health outcomes. It is assumed that there are causal pathways between the upstream, social and the non-medical determinants that directly influence health and survival. For example, living in relatively safe and green neighborhoods may not be affordable for some families. Instead, their contextual conditions force them to live near a busy road or an area where they are

exposed to other environmental stress factors that in turn negatively affect their health. Such associations appear to be intuitive but have often found to be complex and involve multiple intervening features [38, 313]. Similar to the aforementioned example of combining different disability trajectories, when analyzing the impact of societal conditions on health and mortality disparities at population and subpopulation level, we are forced to combine individual life courses.

1.3.1 Life course perspective and health selection

A popular approach to facilitate the analysis of causal pathways is the "life course perspective." Broadly speaking, it is the attempt to investigate continuity and change in people's lives by conceptualizing them as "age-linked transitions" and relating them to the sociohistorical context [87]. In answer to the considerably increasing variability of life courses in modern pluralistic societies, the life course perspective emphasizes the interconnection between different individuals' lives, the relative dependence of life course transitions in a social and historical context, and the importance of agency. Aging is viewed as lifelong process in which individual agents choose and adapt within a framework of opportunities and constraints that they are presented with [25].

The life course perspective provides a way to understand the impact of earlier life experiences on an agent's values and behaviors, which in turn may lead to measurable differences in a variety of outcomes over time. When analyzing health and mortality disparities, there are three crucial concepts related to the life course approach. Epidemiological life course literature highlights the importance of "critical" or "sensitive" periods, like the first years of life. Exposure to unfavorable social and health conditions during these periods is associated with permanent health disadvantages later in life. Childhood adversities have been linked to dramatically increased risks of suffering from serious health conditions, such as ischemic heart diseases and diabetes mellitus [166, 53, 231, 316] and mortality in general [37].

The concepts of cumulative risk and pathways (or trajectories) provide another important theoretical link to explain how late life health and mortality differences emerge. The cumulative risk or disadvantage hypothesis assumes that independent or related sociocultural, economical, and environmental exposure accumulate over the life course in an additive fashion. The pathways approach, on the other hand, refers to the outcome of sequences of exposures to certain risk factors. The exposure to certain unfavorable conditions is expected to increase the probabilities of being exposed to the next disadvantageous conditions in the following life sequence and ultimately determine the disease and mortality risk [37].

A central part of the life course framework is the assumption that individual life courses are embedded in a particular social context and that individuals rely on other individuals' decisions and behaviors. For example, parents usually have an enormous impact on their children (early) life. Among others, they most often provide the context for their children's future health related behavior by the transmission of values about diet, hygiene, risk aversion, and others. Unsurprisingly, it was found that disadvantageous conditions experienced during the parent's own childhood often also affect their children's health [199]. When the life course perspective is applied to more than one generation, it blends with theories of social mobility and selection. With respect to health and survival inequalities, the theory of social selection suggests that health is an important driver for social mobility.

While individuals with good health tend to be presented with greater opportunities for social upward mobility, those with health problems often face more restriction and have only limited access to resources or opportunities [220]. In the case when healthy individuals have better chances for promotion in school and on the job market, one can speak of indirect health selection. The direct form of health selection needs to be distinguished from reverse causality. Although the perceived meaning of the concept of health selection tends to change, it generally refers to a social selection process based on the health distribution at the beginning of life. The allocation to such a distribution is partly dependent on the parent's social position and it is assumed that it affects the probability of developing basic skills to, among others, perform better in school and acquire certain advantageous traits that in turn affect their adult health. Direct health selection not only affect attributes to assent social position but also has a direct impact on later life health. Thus, the understanding of health and survival disparities may require modeling such selection processes [221]. As educational and occupational accomplishments are becoming less dependent on family background [39], such "reversed" selection processes may become more important for the emergence of health and mortality disparities in the future.

Reverse causation may also affect the presumed process of how inequalities in health and mortality emerge. It generally refers to a narrower time band and the effect of health shocks during adult live that have a strong impact on individual's social positions. A classic example of such an event would be a necessary medical treatment that drains the economic resources of an individual or a family [6]. In such case, health inequalities drive the selection into social groups instead of being caused by them. As educational and occupational accomplishments are becoming less dependent on the socioeconomic or family background [39], such "reversed" selection processes may become more important for the emergence of health and

mortality disparities in the future. Considering these parallel processes of indirect health selection, changes in social mobility may explain a large part of the paradox that increased welfare and health investment do not reduce inequality. Inter-generational social mobility may have counteracted such investments through indirectly increasing the homogeneity of social groups such that socioeconomically disadvantaged groups generally become more selected and prone to ill-health [175].

Another approach to explain persistent inequalities in health and survival in modern societies focuses on the cause of death dimension and the above-mentioned theory of epidemiological transition. The stepwise elimination of infectious diseases and the vast improvements in prevention and treatment of cardio-vascular diseases have been found to be highly associated with the recent historical increase in human life expectancy. Some societies have presumably entered an advanced stage of the epidemiological transition where neo-degenerative and non-communicable diseases will become the leading causes of death, and most other societies are assumed to follow the same trajectory [286]. With the increasing impact of these diseases on age specific mortality, active prevention throughout the life course may play a more important role in the future. Other than infectious diseases, it appears that many non-communicable health conditions can be at least postponed through certain kinds of preventive behavior or reduced exposure to harmful environments. The access to such indirect impact factors like knowledge about health-related behaviors or healthier environments is often determined by the social position.

1.3.2 Socioeconomic dimensions of inequalities in health and mortality

Educational attainment is the most widely used proxy for socioeconomic position and thus, often used in the analysis of health and mortality inequalities. The existences of an educational gradient in mortality appears to be one of the most robust and universal findings in the literature on mortality disparities [183, 200, 215, 180]. The use of educational attainment as inequality dimension has several advantageous properties. Generally, education is the most available measure. In contrast to income or occupation, it is relative time independent at adult ages. The onset of poor health later in life does not affect education to the same extent as other measures. Among the proxies for socioeconomic status, it can also be considered as the measure that captures the broadest spectrum of possible determinants of health [69].

The extent to which schooling directly affects health behaviors is heavily debated. Some studies found no or a negative association between higher education and health behavior such as smoking and overeating [11, 149]. On the other hand, it is widely accepted that additional

years of schooling increase the chances for landing better paid and safer jobs, creating a wider social network, and generally having more psycho-social resources, which positively affect later life health outcomes through various indirect pathways [172]. These underlying causes and mediators may be summarized as "health capital," an adaptation of Bourdieu's concept of social capital[66]. Although it is tedious if not impossible to decompose contributions of single causal pathways on later-life health or mortality, the concept appears to be effectively represented by educational attainment [127].

Income and wealth are the two other classic measures for analyzing structural social inequalities and its effects on various life events. The direct effect of individual or household income on determinants of health and mortality, such as access to fresh water and adequate nutrition, appears to be more straightforward than for years of schooling. However, the additional indirect effects, such as the access to opportunities for social participation, are equally difficult to quantify or decompose [185, 228, 65].

There are two general hypotheses describing the potential impact of income on health. The Absolute Income Hypothesis (AIH) assumes an indirect relationship between personal income and average health. The Relative Income Hypothesis (RIH), based on Rogers (1979), on the other hand states that the average population health level is strongly influenced by the income distribution on a society level [246]. While theoretical links have been spun for more than four decades, the relationship between income and health outcomes remains poorly understood. At the root of this is the sensitivity and time variability of income data. Individual level or longitudinal income data is (rightfully) considered sensitive information and hardly ever accessible for larger groups of individuals. When self-reported, perception of social acceptance and recall errors often create large biases. Income data is very sensitive to short- and long-term macroeconomic level changes and the implications of income differences vary by regional, social, and household context.

Similar to educational attainment, aggregated income measures cannot be consistently linked to individual-level health unless we assume the relationship is linear [e.g. 119]. Furthermore, income effects on health can be delayed due to saving mechanisms [154, 276]. Another possible pitfall for modeling the relationship between income and individual health lies in the unobserved re-distribution of individual incomes within households or communities [150]. It is reasonable to assume that individuals who form a household together will share their resources to a certain extent. For example, when two cohabiting individuals have vastly different incomes, individual income effects are very likely overshadowed by the shared

household income [78]. All in all, this makes modeling the association between income and later life health outcomes or mortality very complex.

With the growing accessibility of spatially referenced data, environment and geography have been emphasized as determinants of health inequalities. It has long been hypothesized that inequalities manifest themselves geographically through, for example, processes of urbanization and centralization [144]. Generally, growing inequalities increase the spatial detachment of advantaged from disadvantaged groups. The segregation of rich and poor often affects investment in public goods, security, and other services. In extreme cases, advantaged groups are isolated spatially and economically, not allowing disadvantaged groups to benefit from shared resources [267]. Spatial segregation and unequal access to public services and living space tend to enhance existing inequalities and prevent intergroup relationships and social up-ward mobility. In such contexts, geographic references or postcodes work as a proxy for socioeconomic position and directly and indirectly affect late life health or mortality [118].

1.4 Inequality in mortality and health in the context of Spain and Andalusia

In spite of considerable economic fluctuations, Spain has become a forerunner in terms of longevity and health equality in the last two decades. While the country was ranked among those with the lowest life expectancies in Europe during the mid-20th century due to long periods of political and economic turmoil, mortality has decreased unprecedentedly since the late 1990s [252]. Recent international comparisons have moreover shown that Spain has a highly efficient health care system and ranks among the countries with the lowest mortality differences by socioeconomic group. The low differences have neither been affected by the economic crisis of 2008 nor are achieved at the expense of an expansion of morbidity [238, 176, 178].

Due to the limited access to nationwide, longitudinal or individual level data, it is challenging to investigate trends in structural inequalities and their effects on health and mortality for the entire country. Assessment of changes in health equality tend to be focused on particular periods and areas of Spain. While overall mortality rates were decreasing relatively equally since the 1950s, several studies found persistent geographic inequalities in preventable deaths, a concept referring to causes of death, such as suicide, alcohol related cirrhosis, and road accidents that could theoretically have been averted with the medical intervention or precautious measures [205, 239]. Regional variations in mortality within Spain and within cities have been attributed to differences in these preventable causes of death. Generally, they tend to be higher in the communities of southern Spain as well as in areas which score lower on different deprivation indices [113, 206].

Historically, regional differences and a northeast-south gradient have persisted over more than a century. While Spain's northeastern regions, primarily Catalonia and the Basque Country, have industrialized relatively rapidly during the second half of the 19th century, most southern regions of the country remained largely dependent on agriculture and were never able to establish sufficient home-markets. Such imbalances have naturally led to regionally unequal distributions of wealth and labor, and different levels of vulnerability to downward economic trends [249]. At a population level, these regional inequalities have influenced internal migration patterns throughout the 20th century. While disrupted by the turmoil during the civil war, the majority of the internal migrants during the last century came from rural areas in the south ("rural exodus") and settled in the economically more prosperous regions in the northeast or in the metropolitan area of Madrid. Although migration

appeared to be less economically motivated since the mid-1970s, the metropolitan areas of big cities, also in the South, continued to grow due to immigration while rural areas in the south became further depopulated [247, 258]. Selective emigration and continuous economic instability were also associated with persistently lower levels of life expectancy and quality of life in these southern communities, such as Andalusia [298].

1.4.1 Demographic and economic trends in Andalusia

To contextualize the results from Chapter 2 and 4 below, it is important to understand how national demographic and social trends relate to those in the southern-most autonomous community. Andalusia has about 8.3 million inhabitants (2016), the most populated Spanish autonomous communities [141]. It spans from Portugal along the southern Mediterranean coast and borders Murcia in the south east. Once the port to the new world and one of the most influential and wealthiest regions in Europe (approximately between the 16^{th} and mid-18^{th} century), the fall of the Spanish empire and the loss of its trade monopoly led to a decrease in the relative importance of the region [270].

Since the 19^{th} century, Andalusia has economically fallen behind other regions in Spain. Although it had experienced increased urbanization and demographic developments such as the lagged decline of mortality and fertility, which indicate strong social and economic progress, the economic development could not keep up with other regions in Europe or Spain. This slowly emerging disadvantage can partly be explained by a peculiar "tertiarization" of the regional economy. While other regions experienced periods of rapid industrialization, Andalusia remained mainly dependent on their natural resources, including crops and minerals, and then went through a period of vast increase of the tertiary sector without ever having fully expanded the secondary [15, 190, 233]. Although Andalusia produces half of Spain's mining output and a large share of produce and fruits, the region exhibits high unemployment and at-risk-poverty rates [128, 90]. The latest financial and debt crisis of 2008/09 hit Andalusia exceptionally hard, mainly because vast investments were withdrawn from the previously booming housing market in the region. In the aftermath of the crisis, youth unemployment, evictions, and at-risk poverty skyrocketed [93].

The economic development in the last century naturally had an immense impact on demographic transitions in the region. Andalusia reached the later stages of the first demographic transition comparatively late. Throughout the first half of the 20^{th} century, fertility was among the highest in Spain and differences in life expectancy compared to the national level were mainly contributed by the high mortality of the population aged 50 and younger.

The relatively high mortality in younger people during that time can be traced to various differences when compared to other Spanish regions, like the relatively high importance of labor in the primary sector, low levels of infrastructure development, and macro-political events. Furthermore, Andalusia's population suffered disproportionately under the political persecutions and indiscriminate killings of the Francoist allies and judicial system during and after the Spanish civil war [12, 51, 234].

The dire political and economic situation additionally moved many people to emigrate and look for a better life elsewhere. Between 1950 and 1980, the high emigration from Andalusia was only offset by the extremely high fertility. Beginning in the 1980s, birth rates began to tumble in Andalusia like in many other European countries, often associated with the second demographic transition. Other than elsewhere it was not necessarily linked to a higher influx of women into the labor market nor changing family regimes but more likely to the coinciding democratization in Spain [79]. In terms of longevity and quality of life measures, Andalusia still lags national trends but was able to close the gap within the last decades. Especially child and young adult mortality has decreased and only the population aged 60 and older has not caught up to Spanish national levels of life expectancy [298].

Moreover, Andalusian overall excess mortality is by a larger proportion caused by disadvantages in the male population, which can be confirmed by the relatively large sex gap in life expectancy at birth compared to other autonomous communities and the national level [252]. A closer look at sex-specific labor market contributions and recent migration patterns may offer an explanation for this phenomenon. As above mentioned, Andalusia produces the majority of Spanish crops and has a comparably large mining industry, both related to a cheap and often temporary labor force. In the booming years before 2007, migrants from South America and North Africa came in large numbers to work in agriculture or the construction business. These workers were to a larger extent male and are or were employed in low-paid, demanding jobs. Although a selective migrant population can have an impact on overall mortality differences between men and women, most of the effect is contributed by the native population. Although Andalusian men have been engaged in manual and demanding labor to larger extent than women, it is difficult to make general assumptions about how work related challenges and health risks match those of unpaid labor and housekeeping that especially the women in the older generations were responsible for [117]. The only hint of employment related sex differences can be found in higher male mortality due to road accidents during the 1990s [297]. Larger than usual sex differences in survival are most often rooted in socially accepted behavioral differences that, for example, influence sex-specific smoking and risk

behavior. Here we find evidence for a historical disadvantage for lower educated men in Spain, that slowly appears to decrease due to increasing smoking initiation rates for women [260]. A further male disadvantage was found in mortality due to sexually transmittable diseases, especially due to AIDS during the epidemic in the 1980-90s, which plagued many southern provinces in Spain and led to increased mid-life mortality [298].

Since the onset of the recession, the influx of migrants from overseas has ebbed and migration patterns in Andalusia have shifted. On the one hand, return migration continued while, on the other hand, many young, educated people have left Andalusia (and Spain) in search for better job opportunities [82, 204]. Apart from the long-term effects of selective emigration (sometimes referred to as "brain drain"), it was expected that recession and austerity measures such as the dismantling of the Spanish health care system will negatively affect health, overall mortality, and increase inequality [106]. However, there are known health promoting effects in the immediate aftermath of an economic crisis that potentially offset the negative effects of unemployment and economic insecurity. Among others, reduced economic activity is known to lead to a decreasing in motor vehicle use which in turn leads to a reduction in accident mortality, an association that was also observed during the recent recession [272]. While some researchers assume that the effects of the crisis and austerity on health outcomes will be time-lagged [165], as of today there appears to be little evidence for worsening health or increasing inequalities in survival times in Andalusia or Spain as whole [e.g. 238].

1.5 How to measure inequalities over time?

There is no agreement nor a uniform measure for how to estimate progress in health equality [179]. One can broadly distinguish relative from absolute measures of inequality. In the context of ongoing mortality decline, the estimated changes in inequality may vary depending on the choice of measurement and levels of baseline mortality for different socioeconomic groups. Generally, we assume that baseline levels are higher in lower socioeconomic groups, which in turn leaves them with the most potential for improvement [306]. When changes in mortality equality are, for example, assessed in terms of percentage of original mortality rates (relative scale), improvements require greater reductions in the lower socioeconomic groups. If one calculates changes in absolute inequality (e.g. absolute mortality rate differences), it tends to be easier to achieve larger declines in lower socioeconomic because of their higher baseline levels [178].

The research in this dissertation predominantly discusses group differences in the timing of different life course events, such as the onset of disability. Thus, results are typically presented as relative (e.g. hazard ratios) rather than absolute differences. The survival models that are used in the remainder of this thesis inherently capture the time dimension. The following section aims to provide the reader with a short overview over a few central concepts of survival or event history analysis, which are not explained in the individual research articles below.

The underlying idea of survival analysis is to model the timing of an event of interest based on age specific or longitudinal event rates. In demography, researchers often use these types of models to estimate differences in the occurrence of events across the life course, such as the birth of a child. Originally applied to compare treatment effects over time, survival models have the advantageous property that they allow the incorporation of censored or truncated data. Censoring essentially refers to missing information on the exact survival time and can occur for various reasons. The most common type is right censoring, which often occurs at the right end of the follow-up period when a long-term study is discontinued, or individuals withdraw. Left censoring refers to the situation when the true survival time is less or equal to the observed survival time. It is much less common but appears when the event of interest is the onset of a disease and an individual is tested positive at the first time point. In such a case, one would know that the event has occurred but not exactly when because the left end of the follow-up is unknown. When analysis of time to an event of interested are limited to a certain period or age group, due to study design or data limitations, data tends to be truncated. In contrast to censoring, truncation refers to the exclusion or inclusion of cases

based on artificial boundaries or missing information. In one of the following articles, we estimated the differences in survival probability from age 65 onward. It is natural to assume that some individuals, which would be part of our study population, have already experienced the event, in our case have died, before age 65. Although they would have affected the overall mortality risks and group differences, we were not able to include these cases due to missing information. Thus, our data is left truncated at age 65.

There are a handful of central concepts and functions which I would like to make the reader aware of to facilitate the interpretation of the results below. First, an individual survival time is often given by a T and when we ask if such survival time exceeds a certain point on the time axis, this point would be given by a small t. The survivor function is denoted as $S(t)$ and refers to the probability of not having experienced the event of interest at a given point t. In a life table setting with age as time dimension, this function would be comparable to the l_x curve. Another important concept is the hazard. It can be defined as the instantaneous potential that an event occurs during a time unit under the condition that it has not occurred before [156]. To illustrate, what a survival and a hazard function usually look like when applying the models to adult mortality, the survival probability and the instantaneous hazard are estimated by age and above age 35 for the Andalusian census population of 2001. The two examples were derived from a non-parametric Kaplan-Meier estimation and are depicted in Figure 1.6. The survival curve $S(t)$ on the left-hand side is technically a step function, but generally looks smoother when more observations are included. The right-hand side graph shows the instantaneous hazard function, which fluctuates immensely due to the small time intervals used but increases over age. In some sense the hazard function, denoted as $h(t)$, gives the opposite information than the survivor function, as it describes the instantaneous risk to experience the event at any time t. Mathematically it can be expressed as follows in equation 1.6.

$$h(t) = \lim_{\Delta t - 0} \frac{Pr(t \leq T < t + \Delta t | t \geq T)}{\Delta t} \tag{1.6}$$

, where the hazard at time t is expressed as conditional probability in an infinitesimally small time interval. Based on this function one can assess group differences by applying different assumptions and derive the other functions such as the survivor function, the cumulative hazard $H(t)$, and the distribution function $f(t)$.

Apart from the above mentioned non-parametric Kaplan-Meier technique, there are two main families of statistical models to estimate survival probabilities or hazard ratios, which allow for group comparisons, based on time-to-event data. Parametric models are often

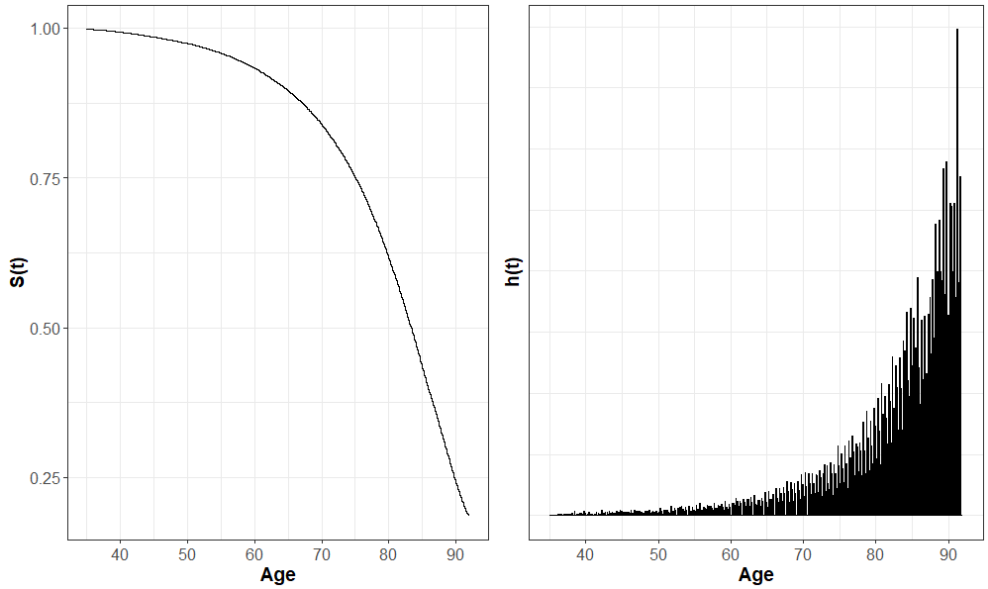


Fig. 1.6 Estimated survival curve and hazard function over time - own estimation with data from the BDLPA

used, when the underlying baseline hazard is known or assumed to be known to follow some kind of distribution. A classic example is that adult mortality over age does approximately follow a Gompertz distribution. When one has additional information about the underlying distribution, the baseline hazard can be specified to estimate the hazard rate at different time points [10]. If the main interest is to model the effects of some set of covariates on the hazard, a semi-parametric approach is often chosen in applied research. The arguably most popular survival model, the Cox proportional hazards model, is a such a semi-parametric model which does not require to make assumptions about the baseline distribution, an advantageous property when the underlying hazard function cannot be specified [197]. As the main interest of this thesis is to outline group differences and potential inequalities in health and survival, versions of these semi-parametric models were applied to the research questions below and are explained in more detail. For more in-depth information about how these models work, how to derive the functions, and how they can be adapted to various scenarios, I would like to refer the reader to the literature cited in this section.

1.6 The essays

As of the day of submission, the three presented research articles in this thesis were at different stages of the publication process. While all three papers are collaborative works of research and were not possible without the help of my co-authors, I (Mathias Voigt) am the

leading or first author in all three cases. The first authorship included the responsibility for the following tasks: formulating the research questions, reviewing literature, choosing the methodology, (in most cases) managing the data, applying and testing the methods, making the results accessible (graphs, tables, text), and concluding the findings. My co-authors supported me with data management tasks that I was not able to do, generally provided access to the data sources, helped me with the understanding of advanced methods, pointed me in the right direction in the process of finding important questions, and reviewed and commented my work to make it better.

The article on urban environment and mortality disparities (chapter 4 of this thesis) was presented at the Annual Meeting of the PAA in Denver, USA, in 2018, the European Population Conference in Brussels in 2018 and published in the journal Population, Place and Space (PSP) in April 2019 and the final, accepted version can be found under the following link:

https://doi.org/10.1002/psp.2239

The paper was published with the help of my co-authors, Dariya Ordanovich from ESRI Spain, Dr. Fransisco Viciana and Rosa Cánovas from IECA, and Dr. Diego Ramrio and Laura Cilek from the IEGD (CCHS-CSIC).

The other two chapters were internally reviewed and found acceptable for submission to a peer-review journal. The work on retirement income and mortality disparities in Andalusia is co-authored by Dr. Fransisco Viciana, Víctor Montañés and Rosa Cánovas from IECA, and Dr. Diego Ramiro from the IEGD (CCHS-CSIC). An earlier version of this paper was presented at the International Population Conference 2017 in Cape Town, South Africa. Chapter 3 on disability pathways was a collaborative work with Prof. Antonio Abellan, Dr. Julio Pérez, and Dr. Diego Ramiro from the IEGD (CCHS-CSIC). It was presented at the Annual Meeting of the PAA in Austin, USA, in 2019 and it is planned to proceed with the editing and submission process within the next few weeks after this thesis was handed in.

Chapter 2

Structural Inequalities and Survival within the Retired Population of Southern Spain

Abstract

Background: Adjustments to the crisis-stricken Spanish public pension system are generally directed at a prolongation of working life and the incorporation of average life expectancy in the pension base formula. When ignoring heterogeneity with regard to survival and well-being, these changes have the potential to enhance existing social inequalities in the retired population.

Method: The impact of retirement income and other socioeconomic features on the survival of retired individuals is estimated by applying stratified Cox Proportional Hazard regression models to a linked register data set which follows up an Andalusian census cohort between 2011 and 2016.

Results: Estimates suggest a higher relative mortality risk for men with the lowest pensions. According to the findings, women's survival is not significantly affected by income inequalities at an individual level. When accounting for household income, results indicate a substantially increased relative mortality risks for females in low income household, which is even exceed by their male counterparts.

Conclusion: Findings confirm the existence of socioeconomic mortality differences in retired population of Andalusia, which substantially vary by sex. Future adaptations of pension systems will need to account for such structural life course differences and its consequences for future population health and longevity.

Contribution: The article extents the literature on income and wealth related inequalities in longevity by analyzing survival time differences of the retired population of Andalusia dependent on pension income and other sociodemographic factors. It furthermore attempts to highlight potential pitfalls of recent adaptations of public pension system in Spain and its consequences for social fairness.

2.1 Introduction

Low birth rates and historically unique improvement in mortality rates across all age groups have led to shrinking and fast aging populations across Europe [237]. After trailing the forerunners in longevity for most of the 20^{th} century, Spain has become the country with the European record life expectancy in 2015 [140]. Such increase of average human life lengths entails considerable challenges for social security systems [56, 126]. While the growing number of welfare and pension recipients increases the economic pressure on the welfare state, the working age population shrinks at high rates. Aging measures like the Old-Age Dependency Ratio (OADR) show that Spain is projected to exceed values of most North and Central European countries by 2050 [261, 209] [1].

The demographic pressure was enhanced by the global financial and debt crisis of 2008, which hit South European economies extraordinarily hard. Austerity policies in answer to the recession then additionally jeopardized the sustainability of social service budgets and public pension funds [152, 168]. As the extension of working lives promises to be a cost-effective adjustment to the population development and would postpone shortages of pension funds to the future [73, 34], the Spanish government implemented adjustable conversion factors in the formula which determines future pension sizes and eligibility ages based on average life expectancy instead of inflation [255, 74].

These changes might reduce the economic pressure on the public funds but extended working lives and reduced benefits also impose a burden on future retirees. From the perspective of social fairness, it is important to ask if such a burden is shared equally. With the implementation of average life expectancy in the formula to determine pension sizes and the eligibility age to full pension in Spain from 2027 onward [cf. 255] social fairness might not be considered. As period life expectancy is calculated from age-specific death rates and population exposure most often observed for single calendar years, it is not designed to capture inequalities within populations and between group heterogeneity with regard to survival will remain unconsidered. Hence, as well-being, health spans, and survival are strongly affect by features of working life course, such as income or occupation type, an increased eligibility age and decreasing the pension levels potentially differently affects social and occupational groups. The analysis of between-group difference has repeatedly shown that, for example,

[1] In the case of Spain the OADR is predicted to reach a value of 0.6 by the year 2050, which can in a pay-as-you-go pension scheme be interpreted as the costs one retired person needs to be covered by less than 2 persons in working ages. The descriptive measure can be mathematically expressed as follows: $D = \frac{N_{65+}}{N_{15-64}}$; where N_{65+} is the population 65 years and older, and N_{15-64} the population in working ages

that the life and health spans of lower educated and less wealthy individuals are shorter, and they are exposed to more variability regarding their ages at death when compared to their more educated or wealthier counterparts [215, 292, 55]. As it is already expected that lower educated and poorer individuals spend less relative time in retirement, an extension of working lives for all would also mean relative longer extension for these groups compared to the better educated and wealthier individuals, who are additionally expected to claim and use more resources to facilitate their healthy life styles to extend their average health spans and further reduce variability compared to other groups[259, 40].

In order to estimate the impact of changing eligibility ages and pension sizes on health and survival in the retired population, it will be necessary to acquire more detailed knowledge about underlying mechanisms which currently shape longevity after onset of retirement. Apart from the aspect of fairness, knowledge about structural mortality differentials by socioeconomic and demographic characteristics is essential for the estimation of expenditures for future care need.

This work aims to contribute to the discussion on structural inequalities in health and mortality at the end of life, by presenting an analysis of survival differences based on socioeconomic features in the retired population of Southern Spain. After the introduction and some remarks on the specific regional context, the linked, register-based longitudinal data infrastructure is described. Third, we fit stratified Cox Proportional Hazard models, to separately estimate the mortality differentials for female and male retirees depending on individual, household income, and various socioeconomic measures between 2011 and 2015. Finally, the findings are presented, discussed, and concluded.

2.2 Background

2.2.1 Structural Social Inequalities and the Association with Health and Mortality

Although most high-income countries provide their citizens with comprehensive and affordable health care, there are persistent inequalities with respect to health and survival [175]. Numerous analyses confirm the existence of survival advantages for better educated and wealthier individuals [176, 183, 200, 215, 180, 177], which appear to particularly affect the mortality risks of younger retirees [263, 311].

Income affects health and survival indirectly, often through the access to opportunities for social participation [185, 228, 65]. There are two major hypotheses explaining the impact of income on health and survival. The Absolute Income Hypothesis (AIH) assumes an indirect relationship between personal income, average health, and ultimately mortality. The Relative Income Hypotheses (RIH) on the other hand relates the average population health level to the shape of the overall income distribution [154, 246]. Following the AIH, the focus of this contribution is on how individual public pension income influences the survival of Andalusian retirees over a time period of five years. As conventional income measures are suspected to be overshadowed by the predictive power of the household wealth [78], a subsequent analysis accounts for the income of the partner and a set of household wealth measures.

2.2.2 Mortality Differentials after the Transition to Retirement

Entering retirement is often associated with a substantial loss of social network resources and can, dependent on the sociocultural context, be perceived as entering a period of cognitive and physical decline [72]. Especially during the transition period, various life domains, such as the daily activities, family relationships and financial well-being, are affected by substantial changes [222]. Depending on the psychological and physical condition, the financial situation, and the availability of a family support network, such changes can trigger a downward spiral regarding the individuals health that ultimately leads to their deaths [171].

Findings on the associations between mortality and the transition to retirement are inconclusive [33, 314, 86, 129, 146, 72], which is partly explained with the varying conceptualization of what retirement means [21]. At an individual level, the decision to retire is furthermore stimulated by various life time events including economic benefits or worsening health, which presents arguably the greatest obstacle for measuring the relationship between the retirement characteristics and survival or morbidity [96, 236]. When someone retires early because of the consequences of an accident or declining overall health, the causation will be reversed. To account for this reversed causation effect, longitudinal health data is needed.

The monthly income of most Spanish retirees is closely related to their life time deposit into social security and the period of contribution to the system. It can therefore be argued that pension income resembles a wealth indicator for retirees and an approximation for the working life trajectory and relative wealth over the life course [22, 240]. Notably, gender differences in various socioeconomic spheres, including retirement benefits, have been comparatively high in Southern Spain, particularly in older cohorts [123, 174]. Generally, women

are more likely to experience a fragmentation of their labor market histories and receive on average lower salaries, most often due to the higher investment in unpaid labor, care giving, and child rearing, [117].

This presumable economical disadvantage does, over all, not lead to worse health or higher mortality. While the sex gap in mortality has widened in Spain and other Southern European countries in the second half of the twentieth century, most probably induced by changes sex-specific developments in risk and health related behaviors [52, 114, 57], the current female survival advantage, measured in years of life expectancy at birth, is 5.52 years for Spain (2014) [2]. The higher female average life expectancy results partly from the lower variability of the ages at death. In other words, their ages at death distribution is more compressed when compared to their male counterparts. In combination with the comparatively low variability regarding the amount contributions to social security funds , public pension income might be a less powerful predictor for female compared to male survival [76, 107]. To what extent gender roles and behavioral difference contribute to this association remains to be determined. On the one hand, women in our study might have experienced a reduced risk to be affected by work related accidents. On the other hand, there is no evidence that women suffer less from economic deprivation or the lack of access to certain goods. Assuming a high degree of solidarity within a household, the indirect effect between income and female survival might be stronger associated with household income.

2.2.3 Mortality Differences in the Context of Spain and Andalusia

A recent comparison of 17 European countries confirms that mortality differentials by socioeconomic characteristics are comparably low in Spain [176]. Spaniards enjoy most of these extra life years free of disability [254] and can fall back on an efficient and comprehensive health care system [67].

Nevertheless, the most recent demographic, social, and economic development might entail various challenges related to structural health inequalities. Kuhlhanova et al. (2012) argue that population health in Spain has not yet caught up to the rapid economic development of the last decades and that wealth and education related risk behavior will have a stronger impact on future mortality [165]. Although it was expected that, in response to the worsening recession, the remarkable mortality improvements of the last decades would slow down,

[2]Calculations based on data downloaded from the Human Mortality Database, University of California, Berkeley (USA), and Max Planck Institute for Demographic Research (Germany). Data downloaded on 26.01.2017 [134]

recently estimated short-term effects on overall survival indicate an even faster reduction of mortality rates since 2008 [88, 168]. If this after-crisis trend is indicative for the future development is heavily discussed. The most plausible assumption seems to be that effects of economic downsizes on the population health are time lagged and will be driven by the worsening health of individuals in low income households [165].

This analysis is focused on Andalusia, the southernmost and with about 8.3 million inhabitants (2016) most populated of the 17 Spanish autonomous communities (AC) [141]. The predominantly rural region has gone through economic bottlenecks and strong selective emigration of young, healthy individuals throughout most of the 20^{th} century. Recent trends show, however, that most economic and population health indicators have been approximated to the Spanish national levels. Analyses of small area differences suggest that only a group of municipalities in the south west of the community still exhibit mortality rates above the Spanish average [207]. Andalusia was heavily affected by the consequences of the recent financial and debt crisis. Unemployment rates and at-risk-of-poverty rates reached a high point in 2015 when they exceeded the Spanish national average by 14.6 percentage points [90] [3].

2.3 Data and Method

2.3.1 Data

In 2002 the Institute of Statistics and Cartography of Andalusia (IECA) began to develop a longitudinal database, the *Base de Datos Longitudinal de Población de Andalucía* (BDLPA)[1]. This register-based, statistical infrastructure allows to follow up death and emigration for a cohort of Andalusians from 2002 up to 2016. Due to the complete linkage to the 2001 Spanish population and housing census, users can access a large variety of sociodemographic variables. The possibility of joint applications with other administrative data sources is a further advantageous feature [137] [4].

To examine socioeconomically driven mortality disparities within the retired population of Andalusia, the BDLPA was linked to individual public pension spells, provided by the Spanish National Institute for Social Security (NISS) for the years 2011 to 2016. Covering all

[3] The percentage of population at risk of poverty and social exclusion in 2015 was estimated to be 43.2% for Andalusia and 28.6% for Spain including Andalusia respectively [90]

[4] The mortality and emigration follow up of a 10% sample of the census based population can be accessed through the protocol on the website www.juntadeandalucia.es

forms of contributory public pensions, the data contains information on various administratively relevant variables, including the average individual pension income, for about 93% of the Andalusian population above age 65 [198] [5]. The contributory pensions are supplemented by a non-contributory form which was implemented to guarantee a minimum pension for individuals with an income below a certain threshold value [299], among them the share of women is disproportionately higher than in the overall population, which is most likely a consequence of aforementioned differences in labor market participation. Consequently, it can be assumed that women will be underrepresented in the contributory pension scheme data. Neither the data on the general tax-funded non-contributory retirement benefits is accessible nor information on private pension schemes. The lack of data on the latter will probably lead to an underrepresentation of the wealthiest individuals.

In order to link the pension spell information to the sociodemographic data, the NISS sample had to be reduced to individuals who were registered in Andalusia during the Spanish population and housing census of 2001. All individuals aged 65, the eligibility age to full pension benefits for most occupational groups in Spain [309], before December 2015 and have received a public social security pension during the observation period are included in the working data set. The minimum age for entering the study was set to age 65 to allow for incorporating individuals who have received a disability pension prior to their 65^{th} birthday.

Since the data has been gained through administrative processes, some incoherencies have occurred during data transmission and the linking process. To prevent biases induced by erroneous data points due to small case numbers at higher ages, the upper age limit was set to age 95. Further precautionary measures are taken for incoherencies occurring because of erroneous retirement information, outliers, and missing data. Individuals who emigrated, or who presumably retired too late (after age 80), as well as about 5400 individuals who resided in extensively large households, indicating an institutionalization, are excluded from the analysis. Due to linking and cleaning processes the initial 1.117.086 individual pension

[5] The pension size is calculated dependent on the years of contribution starting with a conversion factor of 50% for 15 years of contribution going up on a sliding scale to 100% for 35 years or more of contribution. Although exposed to various adaptations, the base rate for public pensions is calculated according to the last 15 years of contribution using the last 24 months as nominal base and previous contributions adapted by the retail price index [300].

The contributory part of the Spanish pension system consists of several subsystems regarding particular occupation types such as coal miners. While the calculation of quotas between the subsystems is reciprocal, the general system still covers about 75% of the contributing members. For more detailed information on the Spanish pension system and attempted changes see [130, 299, 300].

spells are reduced to information on 831.231 individuals. All selected retirees have received a public retirement pension between 2011 and 2015, and resided in Andalusia since 2002 [6].

Accounting for aforementioned gender differences regarding the distribution of income and the importance of the household perspective, subsequently 227,374 married individuals whose spouses are living in the same household in 2011 and have received a contributory public pension during the observation period were extracted to create a proxy household data set.

2.3.2 Variables

Individual and combined household pension income are chosen as principle explanatory variable for the examination of the impact of socioeconomic position on survival. To improve the interpretation of the estimated effects, the continuous income variables are collapsed in four and three categories respectively. These categories are based on quartiles and terciles of the income distributions of the population at risk depicted in Figure 2.1 for individual pension income. To avoid a unidimensional approximation, the highest obtained education degree, the ownership status of the home, and the access to personal vehicles enter the analysis as further explanatory variables and refine the socioeconomic position of individuals. Age, sex, the number of household members, and the area of the main household in square meters were tested as additional socio-demographic variables in earlier models. Neither the number of household members nor the area contributes significantly to the explanation of mortality differences in the population. Both variables were considered in various modeling steps but are excluded from the presented final tables. The birth year is included to account for unobserved timing effects.

For the analysis of couple households' additional information on age differences between the spouses, the highest educational degree as well as the pension information of the partner are incorporated. Household income is approximated by the yearly average sum of all forms of public pension income of both partners. In case one partner has died within the follow-up period, the household income of the remaining partner was approximated by the sum of the individual contributory and the widowhood pension benefits. Apart from the total household income a relative income variable is included, which helps us to distinguish between individuals whose incomes exceed the spouse's income by more than 200 Euro per month from the ones who receive pensions of an equal size or smaller.

[6] For more detailed information on the linking and matching process please contact the corresponding author.

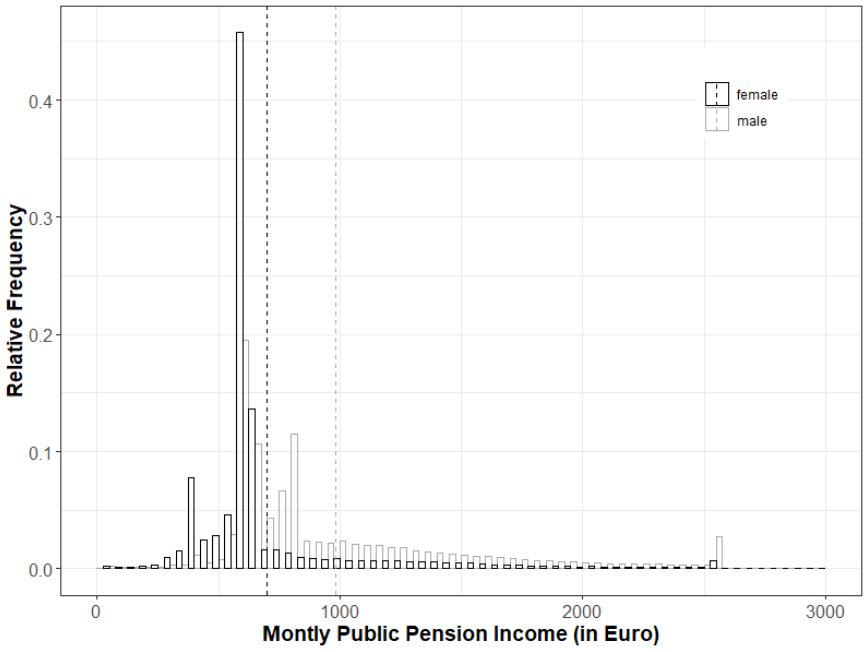


Fig. 2.1 Distribution of Individual Public Pension Sizes by Sex (2011)

2.3.3 Statistical Model

Mortality disparities in are estimated with Cox Proportional Hazard Regressions, which is arguably the most commonly used regression approach to survival data when the focus is on covariate effects [197]. These effects on the individual hazard enter the model multiplicatively as depicted in equation 1 [cf. 156].

$$h(t_j) = h_0(t)\exp(x_j, \beta) \tag{2.1}$$

where $h_0(t)$ is the underlying baseline hazard function at time t and $\exp(x_j\beta)$ the non-negative function of covariates. Model estimates are obtained through the maximization of a partial log likelihood function, a method proposed to maximize the right handed side of the formula with respect to β_x through numerical optimization [10, 282]. Given the setup of our analysis, the model has to account for left truncation [44]. This adjustment affects survival estimates for everybody who has entered the observation period as a pension recipient. Their time under risk before the start date of the study remains unobserved. Hence, the population is selected based on their survival up to 2001.

To account for the different baseline mortality risks and, presumably more important, the differences in working life trajectories of men and women, the model is stratified by sex.

The model is then tested to be statistically different from a model including interaction effects between the set of covariates and the sex variable [156]. Thus, it is assumed that the underlying baselines for men and women differ from each other and justify a separate analysis. To improve readability, it was decided to estimate models separately instead of using a single model with a stratified baseline. The selection and validation of the set of covariates followed a 7 step validation process suggested by Hosmer and Lemeshow (2011) [133]. Model results were additionally compared with the estimates of a flexible parametric proportional hazards model, based on the work of Royston and Parmar (2002)[251], where the baseline failure rate is modeled through restricted cubic splines.

2.4 Results

As the analyses aims to highlight the impact of inequalities by public pension income and other socioeconomic features, structural differences in male and female life courses are accounted for. Two sets of stratified Cox PH models are estimated to obtain relative risks of survival over a period of five years for the two population samples. The estimated hazard ratios and 95% confidence intervals for the best fitting model highlighting the individual public pension effects are shown in Table 2.2. Model results for the survival analysis of married couples are displayed in Table 2.3. Table A.1 and A.2 in the appendix contain estimates and model statistics with pension income and household income respectively as sole predictors of mortality. All models are significantly different from a model without additional covariates. Notably, the number of female pension recipients in the Table 2.2 is substantially smaller than the number of their male counterparts, which is indicative of the aforementioned gender differences in labor market participation and a possible selection bias.

While the effects of individual pension income suggest the existence of a socioeconomic gradient for men, the results for women are inconclusive. For a model with only the household pension income as predictor, the results, presented in Table A.2, even suggest smaller mortality hazards for women in low income households. The hazard for the poorest male group in reference to the highest income category is 82.5 percentage points higher. Estimates between men and women also differ in the best-fitting model where all individuals are included who have received social security pension benefits between 2011 and 2015. The estimated hazard ratios for this group, shown in Table 2.2, suggest the existence of an indirect income effect with respect to survival for the male subpopulations. Compared to the reference group of individuals whose earnings exceed 2000 Euro per month, lower income is associated with a higher relative risk of dying during the observation period. Individuals who are placed in the

lowest income category and earn less than 650 Euro per month are exposed to a more than six percentage points higher mortality risk when compared with the reference group. While the individuals receiving between 1000 and 2000 Euro per month are exposed to an about eleven percentage points higher relative risk, those who earn between 650 and 1000 Euro per month are exposed to a slightly higher relative risk of about 5.8 percentage points. The estimates suggest no detectable effect of individual retirement income on survival for female retirees.

Table 2.1 Descriptive Statistics - Individual Pension Dataset

	Subpopulation	
	Male	Female
N	555,193	276,038
Occured Events	105085	37135
%	18.9%	13.5%
Median Age at Death	81.26	85.23
Average Monthly Pension Income	1025.12 Euro	689.77 Euro
% Less than 650 Euro/Month	29.4%	55.8%
% Incomplete Education or Illiteracy	41.1%	50.5%

The estimates for the indirect effect of the highest educational degree suggest a gradient for both subpopulations. In both cases effect sizes are small but significantly different from the hazard of the individuals with incomplete or no education. For women the lowest relative mortality risk is observed for individuals with a college or university degree, which is about 18 percentage points lower than for those with no or incomplete education. The linkage to the population and household census of 2001 allowed us to include two rather singular measures of household wealth, namely the ownership of motor vehicles and the ownership structure of the household. The estimated effect for these variables suggest a significant direct impact of wealth for men, whereas the ownership structure of the household does only affect women's survival if they own the house or apartment they live in. The access to cars or other motor vehicles has a more pronounced effect on the hazard of males. Men without access are exposed to an almost 22 percentage points higher risk then the reference group.

For women the lowest relative mortality risk can be observed for individuals with a college or university degree, which is about 18 percentage points lower than for their counterparts with

no or incomplete education. The linkage to the population and household census of 2001 allowed us to include two additional measures of household wealth, the ownership of motor vehicles and the tenure status. The estimated effects for these variables suggest a significant direct impact of household wealth for men, whereas women's survival appears only positively affected when they own the house or apartment. The access to cars or other motor vehicles also has a more pronounced effect on the hazard for males where individuals who do not have a care are exposed to an almost 22 percentage points higher risk then those who own a one. While these variables give us an indication for the household wealth, this information is derived from the cross-sectional part of the BDLPA and there is only one data point in 2001.

Besides the income and wealth indicators the civil status at the beginning of the observation period in 2011 is suggested to be a strong predictor for individual level mortality risks. The significantly increased hazard ratios for all non-married individuals indicate a protective and selective effects of marriage. Effect sizes differ slightly between the sexes. Non-married men have an about five percentage point higher relative mortality risk compared to their female counterparts (20.4% vs. 15.5%). The effect of loosing a spouse is also suggest a slightly higher mortality risk for men when compared to women.

The results further confirm that a recognized disability and receiving a disability pension has a strongly significant impact on hazard of dying. Notably, effect sizes are relative similar for males and females. Those who have received a disability pension are exposed to a 42 and respectively 44 percentage point higher relative risk in comparison to those who have never received a disability pension.

Estimated effects and confidence intervals for the two models including results for married couples of 2011, in which both partners have received a public pension before the end of 2015 are displayed in the Table 2.3. This population is more homogeneous than the total population and individuals are generally less likely to die within the follow-up period. Nevertheless, the incorporation of the household income and the other variables related to the socioeconomic position of the spouse contribute decisively to the understanding of mortality differentials in the retired population of Andalusia.

As for the individual pension income, the household or partner income is categorized based on the distribution of the continuous income variable. Individuals who live in a household with a monthly income of more than 2000 Euro are the reference group. Estimated hazard ratios suggest that members of this reference category have significantly better chances to

Table 2.2 Cox PH Model - Individual Pension Income Analysis

	Dependent variable:	
	Hazard Ratios	
	Male	Female
1000-1999 Euro/Month	1.106*** (1.072, 1.139)	1.020 (0.900, 1.139)
650-999 Euro/Month	1.058** (1.024, 1.092)	1.036 (0.915, 1.158)
$<$ 650 Euro/Month	1.063*** (1.029, 1.096)	0.903. (0.784, 1.021)
Reference: $>$ 2000 Euro/month		
Tertiary Educ.	0.964* (0.930, 0.998)	0.822*** (0.749, 0.895)
Secondary Educ.	0.966*** (0.948, 0.985)	0.878*** (0.841, 0.915)
Primary Educ.	0.981* (0.966, 0.996)	0.924*** (0.898, 0.949)
Reference: Illiterate or Incomplete Educ.		
No Car(s) Available	1.219*** (1.206, 1.233)	1.065*** (1.042, 1.087)
Reference: Owns One or More Cars		
Owns House/Aptm.	0.931*** (0.902, 0.959)	0.926** (0.880, 0.972)
Rents House/Aptm.	1.142*** (1.104, 1.179)	1.000 (0.939, 1.061)
Reference: Other Ownership Structure		
Not Married	1.204*** (1.189, 1.219)	1.155*** (1.127, 1.184)
Widowed	1.173*** (1.150, 1.196)	1.123*** (1.092, 1.155)
Reference: Married		
Birth Cohort	1.011*** (1.007, 1.015)	1.021*** (1.013, 1.028)
Received Disability Pension	1.444*** (1.431, 1.457)	1.418*** (1.398, 1.439)
Reference: Never Received Disability Pension		
Observations	555,193	276,038
Wald Test (df = 14)	5,742.060***	1,464.600***
LR Test (df = 14)	5,514.372***	1,444.889***
Note:		*p<0.1; **p<0.05; ***p<0.01

survive longer when compared to the less affluent groups. While the effects for the middle-income category are mixed, retirees of both sexes selected into the lowest household income category are suggested to have a substantially higher risk of dying. The relative mortality risk for women in this group is estimated to be 67.2 percentage points higher when compared with the reference group. Men who live in a household with a combined income below 1000 Euro per month are suggest experiencing a significant 4.8 times higher hazard.

The estimated indirect impact of completed educational degrees also varies strongly by sex. While results for the male subpopulation suggest a decreasing hazard with higher degrees, there is no indication for an effect in females. Other indicators for the socioeconomic position, however, show expected effects. Not owning a car is associated with an increased relative mortality risk, whereas owning a house or apartment is related to a reduced risk. Effect sizes and p-values for the relative income variable which separates the breadwinner individuals from the ones with equal or less monthly incomes than their partners suggest a survival disadvantage for individuals with higher incomes than their partners. Apart from the more economic measures, the effect of disability is noteworthy since it also differs between men and women. While men with disability have a 52.5 percentage point increased relative risk, females with a disability are exposed to a 73.4 percentage point higher hazard ratio when compared to those without a recognized disability.

2.5 Discussion

The recent retirement reforms in Spain is designed to anticipate forthcoming demographic shifts including the ongoing population aging and shrinking. To assure sustainable social security policies, most reforms are directed towards the prolongation of working lives and the reduction of future expenses on retirement pensions. Although, from an economical standpoint, such reforms are almost without alternative [130, 300], particularly the incorporation of average life expectancy into automatically adapting pension formulas presents a treat to future social fairness as it does not account for heterogeneity with respect to socioeconomically rooted health and mortality differences.

As income and wealth are found to determine survival, raising the eligibility age for full public pensions and reducing benefits potentially exacerbate future health inequalities. While those individuals with middle and high incomes can often afford to work longer and enjoy the opportunity to increase their pension benefits, the extension of working life is rather unrealistic for certain manual occupations and service jobs. Furthermore, individuals in low

Table 2.3 Cox PH Models Estimates - Household Pension (Married Couples)

	Dependent variable:	
	Hazard Ratios	
	Male	Female
HH Inc. 1000-2000 Euro/month	1.373*** (1.314, 1.431)	1.087. (0.988, 1.185)
HH Inc. < 1000 Euro/month	4.781*** (4.713, 4.849)	1.672*** (1.547, 1.796)
Reference: HH Inc. > 2000 Euro/month		
Tertiary Educ.	0.816*** (0.707, 0.924)	1.001 (0.796, 1.206)
Secondary Educ.	0.872*** (0.810, 0.934)	1.061 (0.951, 1.172)
Primary Educ.	0.927** (0.878, 0.976)	1.058 (0.979, 1.136)
Reference: Illiterate or Incomplete Educ.		
Received Disability Pension	1.525*** (1.496, 1.555)	1.734*** (1.689, 1.780)
Reference: Never Received Disability Pension		
Birth Cohort	1.050*** (1.040, 1.060)	1.057*** (1.042, 1.073)
Breadwinner	2.108*** (2.076, 2.140)	1.977*** (1.893, 2.060)
Reference: Earns Comparatively Less or the Same		
Tertiary Educ. Partner	1.076 (0.948, 1.203)	0.762** (0.570, 0.955)
Secondary Educ. Partner	0.967 (0.902, 1.032)	0.855** (0.752, 0.958)
Primary Educ. Partner	0.976 (0.928, 1.025)	0.898** (0.820, 0.976)
Reference: Illiterate or Incomplete Educ.		
No Car(s) Available	1.278*** (1.246, 1.310)	1.100*** (1.050, 1.149)
Reference: Owns One or More Cars		
Own House/Aptm.	0.911* (0.840, 0.982)	0.880* (0.769, 0.991)
Rent House/Aptm.	1.140* (1.034, 1.247)	1.041 (0.871, 1.211)
Reference: Other Ownership Structure		
Lives only with Partner	0.919*** (0.888, 0.949)	1.003*** (0.955, 1.051)
Reference: More Household Members		
Observations	113,690	113,684
Wald Test (df = 21)	7,569.420***	1,644.360***
LR Test (df = 21)	6,325.701***	1,613.716***
Note:		*p<0.1; **p<0.05; ***p<0.01

income households face more barriers with regard to private pension provisions which, in turn, makes them more vulnerable to reductions of pension sizes [255].

As simulations considering varying eligibility ages for full public pensions predict increasing inequalities in the retired population of Spain [73], this work aims to contribute to the discussion by estimating effects of pension sizes on individual level mortality in Andalusia. Identifying disadvantaged groups within the retired population of Spain, a country which offers efficient public health care to their citizens, can raise awareness for the potentially unanticipated consequences of the recent adaptations to the public pension system [54, 67].

Considering the findings from two different sample populations, results suggest a substantial effect of pension income on the survival of male retirees. Conclusions drawn from the effects of individual and household level retirement pension, are based on the assumption that public pension income is strongly associated with the lifetime earnings and the fragmentation of the labor market history. The results of the household pension analysis suggest that males in low income households face a strikingly 5 times higher mortality risk when compared to the richest third of the population. While this effect is enhanced by the selection of couples who receive retirement benefits, it gives an indication for how socioeconomic disadvantages can accumulate across the life course and culminate in extreme inequality regarding the ages at death. Although the effects are less conclusive, mortality risk also appear affected by individual income measures. Remarkably, when applying individual income, women appear to be immune to the negative effects. Only females who live in a household with a monthly income of less than 1000 Euro have an elevated relative mortality risk when compared to the highest income group. As numerous indirect factors might influence individual mortality risks, the presumable female immunity to income differences is likely caused by the non-random exclusion of women who have not payed into the social security pension funds [76]. The ones who have contributed often experienced disrupted and short-term labor market histories. Individual level career choices of these generations of Andalusian women were at least partly influenced by social norms, which among others promoted a male breadwinner model as preferred family formation. Hence, the socioeconomic position of these females is only poorly reflected by individual pension income [22].

Apart from the gender differences regarding the labor market participation and the consequential dissimilarities in the income distribution, at a descriptive level the ages at death are much more compressed for women than for men. Life span disparity measures such as the years of life expectancy lost to death ($e^{\dagger}$) confirm that female age at death tend to be

more compressed. They confirm that overall women seem to die more equally with respect to the age at death. While we know that the variations between men and women who die prematurely are mainly affected by differences in cardiovascular mortality in Spain [109].

A time lag between an income situation and the effect on health might have proposed a further tread to validity [154, 276]. Even if the direct income hypothesis can be partially confirmed with our results, estimates can neither be unequivocally related to the latest economy crisis nor pension level reductions. Moreover, the duration individuals compared in the sample population have received a public pension varies substantially. Hence, if someone received a public pension over a long period of time she presumably respond differently to changes in the pension size compared to someone who just entered retirement.

Several limitations are related to access and administrative structure of the data. The time lag between the census and the pension spell data restrict the analysis to time independent variables like sex or the highest education degree. Other potentially useful information on the household and individual level could not be incorporated because of the ten year lag. Changes in the ownership structure of the house and the access to motor vehicles were not assumed to be extensive, although interpretations could be affected.

The time lag also affects the representativeness of the sample population whose members had by definition to be registered in Andalusia in 2001 and received a public pension before the end of 2015. Migrants and return migrants who entered or reentered Andalusia after 2001 are not considered. The popularity of the coast of Southern Spain as retirement domicile for retirees from other European countries leads to the assumption that a large part of the European migrant population is not be covered by our data [163, 214]. Further limitations are related to initial difficulties to disentangle the disability pensions from retirement benefits before age 65. Whereas disability pension is on paper automatically turned into a retirement pension after age 65, the situation remains unclear for younger age groups. Thus, it was not possible to estimate the effects of early retirement on later-life mortality. The data driven confinement to the occurrence of death as the only possible outcome was further limiting element particularly affecting external validity. This work attempts to quantify mortality differences based on existing structural inequalities within the retired population of Andalusia and draws inferences for potential long-term consequences of the latest reform of the Spanish public pension system. To our knowledge it was the first time that someone had used individual-level retirement spells linked with a mortality follow-up to analyze the association between pension income and survival in the context of Spain. Although the usage

of administrative information was limiting, the substantial wealth related inequalities that we find to be associated with old age survival are alarming. The findings support the hypothesis that incorporating average life expectancy to assess how much working lives need to be extended or pension level decreased does neglect persistent health inequalities. Even if the analysis contributes more to a better understanding of mortality differentials in the today's elderly population of Andalusia, inferences can be drawn for the future development.

Considering that the working age population today faces a strong insider-outsider labor market, it will be necessary to account for the looming increase in old-age poverty in Spain [29]. Even if these younger generations appear to be healthier and better educated than the boomer generation [158], their chances to accumulate sufficient public pension benefits or private provisions is severely limited by the hostile labor market situation. Regarding the situation in Andalusia it will also be important to investigate the effects of increased participation of women in the labor market. Will the health of future female Andalusians vary more with regard to age at death or the age at onset of disability? Or is another scenario possible in which men benefit in terms of survival?

Generally, it will be important to account for heterogeneity with respect to life expectancy. Given the remarkable public health interventions in European countries, the focus of future analysis of this relationship should be rather on the quality of life, well-being, and the onset of morbidity and disability. Remarkable gains in life expectancy were possible because the majority of the population benefited from the comprehensive implementation of social security benefits, public pensions, and access to health care. If measures who are directed to assure sustainability of pension and social security funds ignore the existence of frail or disadvantaged groups, inequalities in health and ultimately life spans will be further manifested and enhanced.

Chapter 3

Pathways through Dependency and the Effects on Old Age Mortality

Abstract

Background: Population-wide studies find persistent health inequalities when analyzing the effects of one's social position on the age at onset of old-age disability or age-specific risks of dying. It is however less clear how socioeconomic measures affect the risk of death after the onset of disability, which might be related to the lack of a universal definition or common classifications of disability.

Method: A spell-length sensitive matching algorithm is applied to identify commonly shared trajectories after onset of disability after age 50, based on information about the first three successive onsets of limitations in activities of daily life after age 50. Mortality hazards are then estimated with a Gompertz proportional hazards survival model. Effects of different socioeconomic and physical measures are compared within trajectory groups and the healthy population by taking advantage of newly linked data.

Results: Results suggest higher mortality risks for the less affluent and less educated individuals in the healthy population while the effect of such social inequalities appear to become neglectable after onset of old-age disability. Differences in mortality after onset of disability are mainly associated with the lack of daily activity and the time spent with disability.

Conclusion: The study suggests a substantial reduction of socioeconomic mortality differences after onset of old-age disability that appears to be independent of nature or trajectory type of disability.

3.1 Introduction

The exceptional increase of average human life lengths entails the question whether or not the additional years are typically spent in good health [126, 237, 191]. Although prevalence rates for many aging related chronic conditions and disabilities have increased in most developed societies due to ongoing population aging, functionality and quality of life after onset of disability have generally improved [189, 105, 63]. The relatively small improvements in the very old ages and increasing health inequalities in younger age groups, on the other hand, have raised concerns about less optimistic future scenarios which might bring high levels of inequality and drastically increasing health care costs [47, 304, 19, 265].

There is strong evidence suggesting that the onset of morbidity and age-specific mortality are closely associated with income, wealth, education level, gender, ethnicity, and occupation. Better educated and wealthier individuals experience a later onset of old-age morbidity and survive on average longer when compared with their less wealthy and lower educated counterparts [176, 183, 200, 215, 180, 177]. It is assumed that such disparities ultimately derive from structurally restricted access to key determinants of health such as fair pay [36]. While the association between socioeconomic measures, late-life morbidity, and age-specific mortality is well-established, surprisingly little is known about how mortality is affected by the socioeconomic position after onset of disability and in the light of continuous medical and social progress [121, 310]

The objective of this study is to identify and characterize different health trajectories after the onset of old age disability and then analyze the effect of socioeconomic context variables on old-age mortality within groups with shared trajectories. The aim is further to extend the research on the well-known relationship between socioeconomic position, late life health, and mortality by estimating time-to-death differences in dependence of different disability pathways and socioeconomic features. The effects are then compared to mortality differences within the disability free population.

We hypothesize that effects of socioeconomic position are generally less relevant for the survival after onset of disability. It is also assumed that individuals who have experienced a comparably early onset of disability may be more resilient and affected by their socioeconomic context when compared to those who experience a late onset. Due to the persistent economic disadvantages women in this study have generally faced during their working ages, it is further expected that socioeconomic characteristics will be more dispersed for men and, thus, lead to larger effects on male survival.

3.1.1 The assessment of health and its association with mortality

The postponement of the onset of chronic conditions or disability in a population closer to the average age of death is commonly referred to as compression of morbidity. Within the framework introduced by James Fries and colleagues in the 1980s, it is assumed that mortality asymptotically approaches a maximum age while the average onset of morbidity is further delayed to higher ages due to medical and social progress [102, 103]. Although it was criticized for its sole focus on chronic conditions and the assumption of linear decline, which theoretically disallows for recovery, Fries' framework and the adaptations to it, like dynamic equilibrium [cf. 182], laid the ground for today's assessment of whether or not additional life years are more likely to be spent free of disability [24].

Most uncertainties about how mortality and health are associated in future population health scenarios are rooted in difficulties to define and measure health and disability. While death is a clearly defined concept and mortality information is routinely collected by statistical offices, it is far more challenging to assess complex concepts like health. When analyzing old-age morbidity trends at population level, researchers are often forced to combine individual life courses with vastly different health trajectories [112].

Even if health is solely defined as absence of disease or chronic condition, when diagnosed, many health conditions can range from mild effects to being incapacitating. Thus, to distinguish a healthy from a sick person becomes increasingly tedious the larger the variety of considered conditions [35]. To reduce complexity and instead of assessing single medical conditions, public health researchers commonly fall back on examining performance in activities of daily living (ADL) and instrumental activities of daily living (IADL) as proxies for morbidity [138, 155, 269]. Most commonly, those ADLs include bathing, dressing, eating, and toileting while IADL refer to periodic routines often including meal preparation, shopping, housekeeping, laundering, using the phone, managing medications, managing money, and using transportation [167].

3.1.2 Shared pathways

The substantial amount of possible health conditions, variations by onset ages, effects of comorbidity, the need for longitudinal data, and generally different degrees of activity limitation prevent a uniform assessment of health and make a classification of disability trajectories a challenging task [111, 101]. When examining such trajectories or pathways, the increasing availability of linked databases, often traditional, quantitative sources of

information on disability such as health surveys or public registers which are linked to some kind of mortality data, have added a new dimension that allows researcher to investigate temporal and causal relationships between disability and old-age mortality [232, 112]. Such a newly linked data source allows for identifying disabled individuals and their household members in Spain and following them up to death or censorship in 2018. In the attempt to classify different pathways and then analyze the impacts on old-age mortality by each path, the remainder of this work is structured as follows. First, the linked data sources and the study population is described. Second, it is explained how the definition of disability is derived and how a matching algorithm can help to identify shared patterns based on the first three episodes of disability. Third, the statistical model for the survival analysis is explained and results are presented. Last, we conclude findings, limitations, and implications on future research.

3.2 Methods and Materials

3.2.1 Data

Initiated by the National Institute of Statistics (INE) and conducted in 2008, the National Survey on Disability, Personal Autonomy, and Dependency (EDAD, Spanish: *Encuesta sobre Discapacidad, Autonomía personal y Situaciones de Dependencia*) is one of the largest national survey studies related to the topics of health, disability, and care in Europe. Stratified samples for this cross-sectional study were drawn in a two-stage process. The initial sampling stage, based on 3,843 Spanish census tracts, was adjusted in size to allow for more variation by single autonomous region. In a second step 96,075 households were drawn from the selected census sections, which led to a final sample of 258,187 individuals between the ages 0 and 104. The data was collected through face-to-face interviews between November 2007 and February 2008 (overall response rate 97%). When a household member aged 6 or older was identified as disabled according to a prior definition, the person or, if applicable, her care giver were invited to answer an individual questionnaire directed at health related topics and personal experiences with disability. A second survey stage was aimed at sampling individuals from health institutions and nursing homes [139].

In collaboration with the INE department for socio-demographic statistics the great majority of individuals surveyed in the EDAD study was linked to administratively collected,

longitudinal mortality and exposure data for the period between 2008 and 2018 [1]. Mortality records were extracted from annually updated statistics of natural population movements (MNP, Spanish: *Estadística de defunciones. Movimiento natural de la población*) and linked via personal identifier. Exposure rates are based on yearly updated data from the population register (Spanish: *Padrón*) and allowed capturing migration movements.

In contrast to previous, nationally representative surveys on disability, EDAD is oriented towards the assessment of how individuals with disabilities handle their daily life and how they are integrated in everyday social life [3]. Building on the WHO International Classification of Functioning, Disability and Health (ICF), the questionnaires are directed at the efficiency of personal and technical support and the importance of social participation. Instead of being examined medically, the respondent is asked about limitations in activities of daily living. Where possible, these limitations are traced back to an underlying disease, chronic conditions, or injury [4].

3.2.2 Defining disability

A person is classified as disabled if she has experienced difficulties in performing at least one of the following 13 activities. The list includes 10 activities of daily living (ADL), namely body position changes, walking indoors, bathing, basic hygiene, urination (bladder control), toileting (bowel control), dressing and eating. The ADLs are complemented by five instrumental activities of daily living (IADLs), which include shopping, preparing meals, walking outdoors, using public transport, and being able to do housework. Figure 3.1 gives an overview about the selected items.

3.2.3 Characteristics of the Study Population

To observe only disability patterns representative of those commonly experienced by older people, the sample is restricted to all non-institutionalized individuals who experienced the onset of disability after age 50. Furthermore, the disability onset had to occur at least two years before the survey years to allow for the necessary variation in the individual trajectories and prevent misclassifications. The final sample contained 7855 cases (5450 female, 2405 male).

[1] About 1000 individuals who part-took in the survey could not be linked and followed up due to various reasons including unregistered emigration

List of ADLs and IADLs to define the disability status

Activities of daily living (ADL)

- Body position Changes/Getting in or out of bed
- Walking indoor
- Walking outdoor
- Using public transport
- Bathing/showering
- Basic hygiene
- Urination (bladder control)
- Toileting (bowel control)
- Dressing
- Eating

Instrumental Activities of Daily Living (IADL)

- Shopping
- Preparing meals
- Housework

Fig. 3.1 Activities included in the Assessment of Disability (ADL, IADL)

The selected individuals are substantially older than the total population aged 50 or older, which is composed of all individuals in the surveyed households. Their median age in 2008 is 77 for men and 78 for women compared to a median age of 63 for both sexes in the total population. The average onset age of disability after 50 is 69 years for both sexes. Further descriptive statistics comparing the two groups are provided indicate that the older age structure of the disabled population also drives the ages at death to be higher.

The distribution of sociodemographic variables differs more for females than for males when comparing the two samples. The majority of men in both populations is married (76.3% in the disabled population, and 83.3% in the population without disability). While most female participants in the disability-free population are also married (67.7%), the majority of the women with disability is widowed (49.1% compared to 42.2% in marriage). Due to the relatively late expansion of education in Spain, a relatively large proportion of the population does not have completed a formal degree. In the disabled population 59.5% of the females and 48.8% of the males have, for example, not finished school.

Income is measured as monthly gross income per household consumption unit (CU), which are derived from the household composition. Weights are assigned to every household member based on the OECD modified equivalence scale [2] [97]. The distribution of monthly income per CU indicates that more than half of the individuals with disability live from less than 750 Euro a month (median income for both sexes). When income is self-reported, as in the EDAD survey, it is likely to be affected by different kinds of reporting biases such as that people with very low or very high incomes often prefer not to report their income at all [164].

[2]Consumption units are calculated for each household. The household head or first adult is assigned a weight of 1 CU, every other adult is assigned a weight of 0.5. All household members aged 14 or younger are assigned a weight of 0.3.

Furthermore, the variable does not allow for capturing other private assets or other hidden forms of income.

3.2.4 Identifying Shared Pathways with Sequence Analysis and Weighted Clustering

As effects of limitations in activities of daily life on survival most certainly vary by levels of severity, the timing of the onset, and the duration spent with disability, this work aims to identify shared trajectories in people with disability that can then be further examined and compared. To analyze dissimilarities between individual life courses after onset of disability and to identify underlying shared pathways, different data mining techniques can be applied, which compare and group individual state sequences or life course trajectories with minimal prior assumptions [31, 244, 2].

Therefore, the age at onset and the number of newly occurred limitations at the first three successive episodes of disability is used to create state-sequence formatted dataset incorporating all individuals with at least one disability before 2006. In other words, the life course of individuals with disability is followed up from age 50 and the level of disability at every single age is assessed up to when the individual died or is right-censored. Naturally, the big majority of the subjects is assigned the state "disability free" at age 50. Then over time, individuals experience different events that successively either leave them with a mild disability (M), defined as one additional limitation per new onset, with the exception of problems with food or drink intake, or a severe disability (S) referring to the simultaneous onset of at least two different limitations or the onset of problems with eating and drinking. Individual state sequences are formed under the assumption that former disabilities cannot be cured. In other words, an individual can for example never go back to the state "disability free" after it had experienced an event. Figure 3.2 shows the relative distribution of possible states, which range from one mild event (D) to three severe onsets (SSS), by single age between the ages 50 and 100.

A spell length sensitive optimal matching algorithm (OM) is used to quantify levels of mismatch between each individual sequence of yearly updated disability information [cf. 125]. As measure of dissimilarities, this algorithm compares single trajectories with all other cases in a dissimilarity matrix and minimizes the residual variance between the state sequences [42, 275]. The OM distances proposed by this measure are sensitive to differences in duration and the order of successive states. This has the advantage that it better

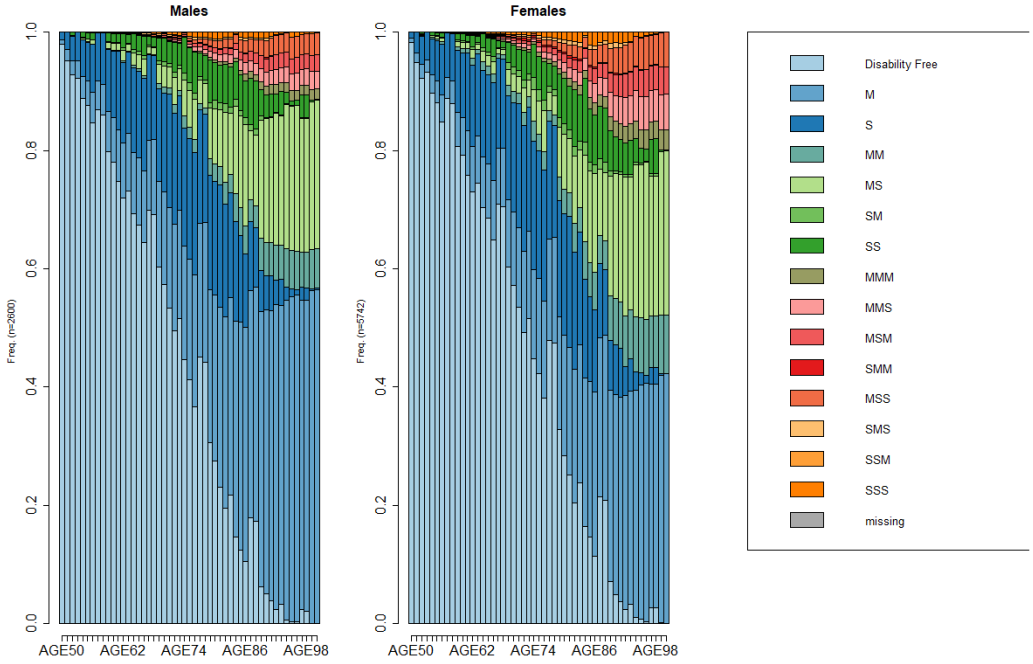


Fig. 3.2 Sequence Distribution by Age and Sex with legend

distinguishes between long-term continuity and rapid changes of states [274].

Based on the dissimilarity matrix, several potential clustering options can be extracted. They are assessed and compared with the Average Silhouette Width (ASW), which is a within-group coherence measure that accounts for sequence dissimilarities within each group [275]. The analysis of dissimilarities and the clustering was conducted using the tools provided in the *TraMineR* and *WeightedCluster* R-packages [cf. 104, 273].

3.2.5 Regression Model

Mortality hazards are then estimated for the population classified as disabled using a Gompertz proportional hazards survival model. The Gompertz model assumes that the logarithm of mortality risks grows linearly with increasing age. It has been shown repeatedly that it models human adult mortality accurately with only two unknown parameters [58]. The models are stratified by previously identified disability trajectory and sex. The hazard function of the Gompertz model can mathematically be expressed as follows.

$$h(t) = a * \exp^{(bx)} \tag{3.1}$$

, where a and $b > 0$ are the rate and shape parameter of the model. Regression coefficients, which are estimated through Maximum Likelihood Estimation (MLE), are asymptotically normal distributed. When comparing mortality between different subpopulations, proportional hazards approaches do not allow for capturing that variances in life lengths are inversely related to average life expectancy. In other words, the proportional hazards assumption imposes the same variance in length of life on all observed subgroups, which is often the opposite to what is observed in reality [284]. While this behavior often requires to adapt the model with an additional frailty term, only one model (disability-free male population) was found to be affected. As the proportional hazards assumption was found to hold for the large majority of variables in all models, it was refrained to include an additional individual frailty term into the models [cf. 295].

As individuals enter the study at different ages, the time under risk of dying before the start year of the study remains unobserved and it is therefore necessary to account for left truncation [44]. In other words, individuals are selected based on their survival upon two years before the EDAD survey. The use of person years as the time scale in our models additionally allowed us to measure age-specific mortality differences rather than survival over calendar time [48].

3.3 Results

The results from the ASW suggest that three clusters are optimal to represent the variation in individual disability trajectories in the data. These three pathways are labeled after a visual assessment of the relative state distributions at different ages. Generally, we observe the same three pathway groups for both sexes with a slightly earlier onset of disability for females. The relative distribution of states by pathway and age are shown in Figure 3.3. The sex-specific pathway distribution plots can be found in the appendix B (Figure B.2 for females and Figure B.1 for males).

The relatively small *mild disability* group consist of individuals with one mild onset (defined as 1 disability at time of the event) who did not have experienced a second event within a period of at least two years. The exception is a small proportion of individuals who experienced further episodes of disability at higher ages. Individuals in the other two groups have, on contrary, either experienced a severe onset of disability, defined as 2 or more limitations occurring at the same age, or have suffered from multiple successive events (disability onsets) that have worsened the severity of their disability over time. These two

groups can be distinguished by the timing of first onset and are therefore labeled *early severe* and *late severe*. Although the relative distribution plots suggest that individuals in these two groups experience multiple onsets of disability at higher ages, on average they spent most of there time under observation in the state S which refers to the occurrence of a single, severe onset. In fact, only 40% of the individuals in the *early severe* group have experiences a second event within the observation period. Within the *late severe* group, men appear to be less likely (51%) to experience from a second onset than women (60%).

In Figure 3.4 the relative distribution of individuals is presented by sex and pathway

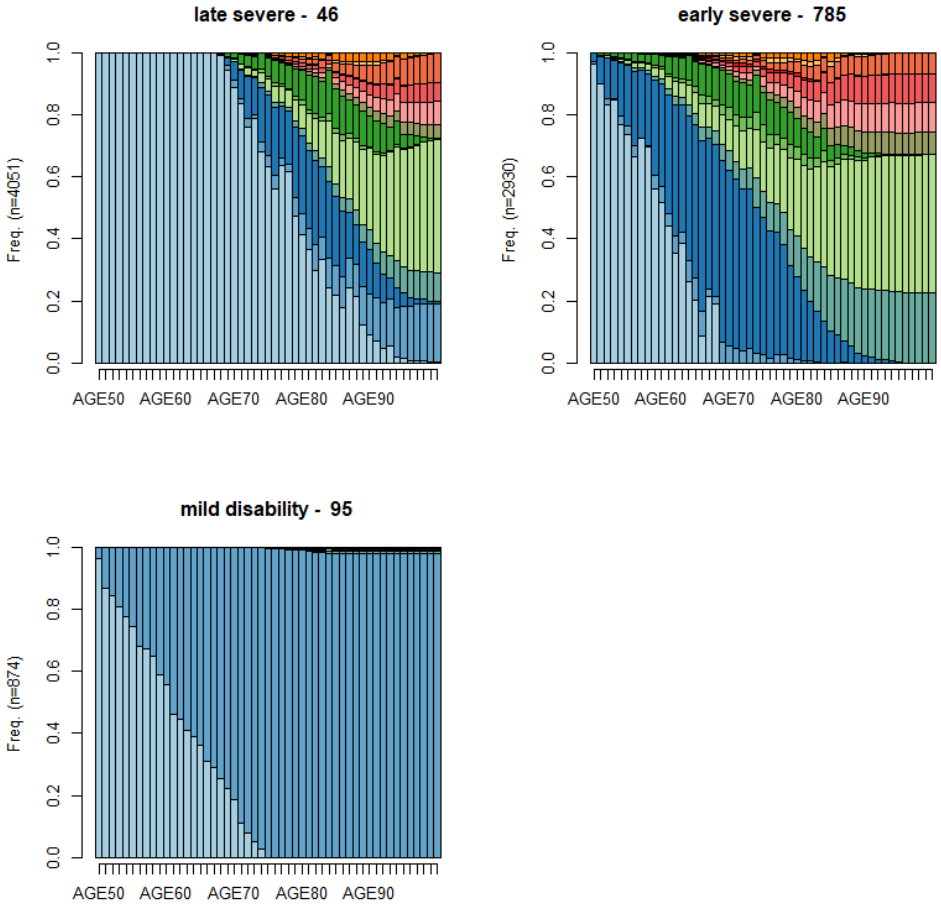


Fig. 3.3 Relative Distribution of State Sequences by Pathway and Age

group. The colors can further be used for a first descriptive assessment of mortality by group and sex; the darker tones represent the proportion of survivors at the end of the follow-up period and the lighter the ones who have died. The graphs suggest that the *late severe* group

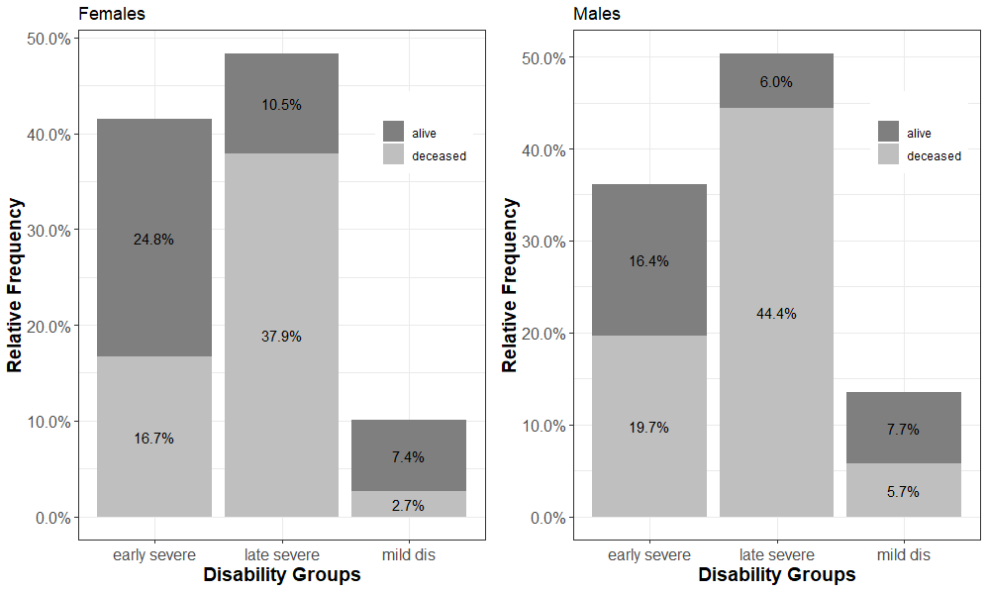


Fig. 3.4 Relative Distribution of the Event of Interest Depending on the Assigned Disability Trajectory

experiences the highest fatality rates in the follow-up time while most individuals assigned to the *mild* trajectory survive until the end of the observation period. Across all trajectory groups, females exhibit a higher probability of surviving up to the end of the study period than their male counterparts. While not only a substantially larger number of females was selected in our study, there is also a lower percentage of male survivors in all groups. The Kaplan Meier survival plots in Figure 3.5 confirm these differences when comparing age-specific mortality. The survival probabilities of disabled individuals aged 50 and older are compared here to individuals in the surveyed households who do not suffer from a disability. While age-specific survival probabilities of those without disability follow age-mortality patterns that are commonly observed in modern developed societies (low mortality up to higher ages and low variability at ages of death), disabled people follow very different patterns. Survival probabilities at each age after 50 are lower than in the reference population and differences (area between the curves) appear to increase over time before they converge at higher ages.

The differences between males and females are substantial. While the curve for females with disability appears to have a similar shape than the references populations with lower levels of survival, the disabled male population experiences an almost linear decline of survival probabilities and vastly lower levels of survival when compared to those without disability. It can be assumed that the survival of the subpopulation with disability might be even worse since cases who die shortly after onset of severe disability were not be observed. Furthermore,

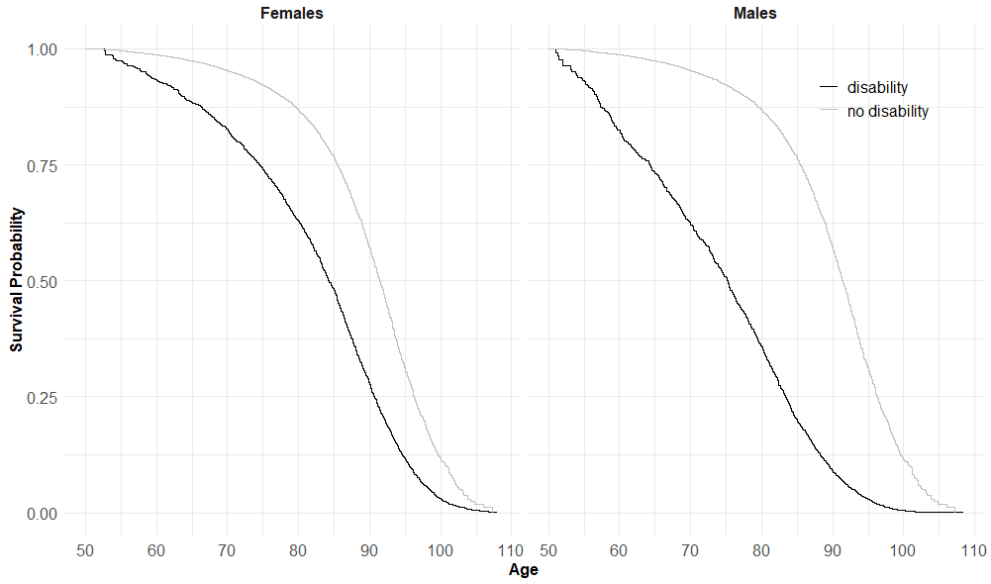


Fig. 3.5 Kaplan Meier Estimates for Survival Probability for both Sexes by Health in 2008

it was not possible to observe onsets of disability after 2008.

When comparing the mortality between the pathway groups, it is important to account for differences in the age distribution within each group. The *late severe* group is on average substantially older than the two others which indicates a higher risk of selection upon survival. This issue is accounted for by comparing age-specific survival probabilities, as for example depicted for estimated survival probabilities in Figure 3.6. These Kaplan-Meier survival curves indicate that the *mild disability* group is exposed to higher chance of survival up to the oldest ages. The shape of the slopes further suggest that an *early severe* onset is most lethal in younger ages while the *late severe* groups slope crosses at it later ages. At the ages 95 and older the curves become less indicative due to the small case numbers.

3.3.1 Impacts on Mortality within Pathway Groups

The descriptive assessment of onset, severity, and underlying baseline distributions suggest that between group differences may cancel each other out and influence the effects physical or socioeconomic features. Thus, in the regression models individuals who fall into a trajectory group are solely compared to the other individuals in that same group. While comparability between groups is intended, therefore the use of the same set of covariates, the comparison of effect sizes between disability trajectories might not be meaningful due to different underlying baseline hazard distributions.

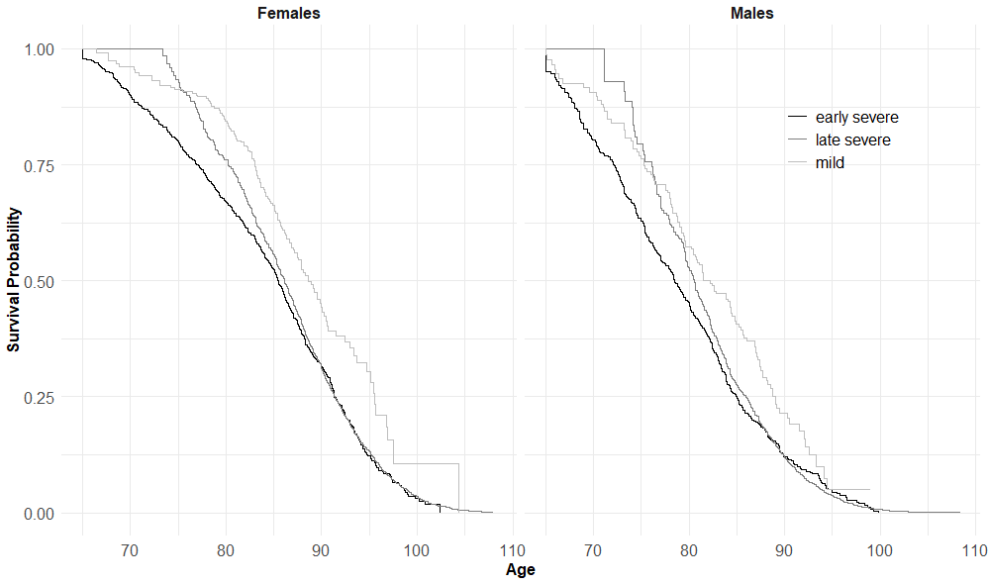


Fig. 3.6 Kaplan Meier Estimates for Survival Probabilities for Both Sexes by Assigned Disability Trajectory

There are tables with six sets of nested Gompertz PH regression. These models are stratified by disability pathway and sex. Table 3.1 and 3.2 contain estimated hazard ratios, confidence intervals, and overall model statistics for a survival analysis of individuals who are selected into the *early severe* pathway group. All models are significantly different from the Null model without additional covariates. Goodness of fit tests suggest that the use of additional socioeconomic and household composition variables does not improve the model fit for the males while there are a slightly significant improvements for females.

The most robust effect in all models is the impact of daily activity. For the women in the *early severe* trajectory estimates suggest that mortality hazards for those who lack daily activity are between 42.9 and 50.6 percentage point higher. For males the values vary between 51.8 and 58.4 percentage points by model. While the exposure to other diseases or chronic conditions does not appear to affect the hazard of dying for neither females nor males, the time between onset of disability and 2006 is suggested to have an effect. The 40 to 60 percentage points higher risk of the group with shortest duration in disability (females between 40.9 and 42.1; males between 59.0 and 61.0) relative to those who have lived the longest with disability within this group can be interpreted in different ways. First, it can be assumed that a severe onset of disability is related to a timely close death and those who have experienced an early onset are exposed to higher mortality risks right after the event

compared to those who have experienced the onset a long time ago. In other words, it can be assumed that individuals go through a high risk period shortly after onset and then adapt to a life with disability. Like for the observed indirect effects of widowhood, the fatality of a severe disability may be strongest directly after onset and then levels off.

Such an adaptation or frailty effect can also be found in the male population in the *late severe* pathway group, shown in Table 3.4. The model results for females in the *late severe* group, represented in Table 3.3 are not conclusive. For both groups the estimates also suggest a strongly negative effect of lack of daily activity. Effect sizes are very similar for both sexes and vary between 50.1 and 52.7 percentage point higher risks between the different models. In contrast to the *early severe* group, suffering from additional diseases and chronic conditions is found to slightly increase mortality hazards for both sexes suggesting that aging related processes limit the individuals ability to respond to the challenges that come with the onset of disability. The estimates suggest a 16 to 17 percentage point increased hazard for males with comorbidity and an about 18 percentage point higher hazard for women in the same category when compared to their counterparts without further conditions.

The results for Gompertz models regarding the *mild disability* trajectory group can be found in the Tables B.1 for females and B.2 for males respectively. The estimates suggest that none of the included variables does significantly affect the mortality hazards for these two subpopulations, which might be due to the relatively small case numbers. To improve readability these model results are presented in the appendix with the other additional material.

To reaffirm earlier statements about the relative importance of socioeconomic position on survival disparities at older ages, mortality hazards are also estimated for individuals aged 50 and older who are free of disability. The estimates for the Gompertz PH models are presented in the Tables 3.5 for females and Table 3.6 for males respectively. As the EDAD survey was aimed to assess life circumstances in people with disability, there is only a small amount of information on the household members without disability. Information about income categories, educational attainment, and civil status are available and included in these models. Household size entered as additional control variable. All models exhibit a significantly improved fit when compared to the less complex, nested models or the model without covariates. The estimates suggest that the socioeconomic position plays a much larger role in the disability-free population.

Table 3.1 Gompertz PH Regression Models - Females with Early Severe Disability Trajectory

	Dependent variable: Relative risk of dying		
	(1)	(2)	(3)
No daily activity	1.506***	1.459***	1.429***
	(1.286, 1.762)	(1.245, 1.710)	(1.218, 1.677)
Reference: Daily Active			
Suffers from multiple diseases	1.157	1.144	1.132
	(0.956, 1.400)	(0.945, 1.384)	(0.934, 1.371)
Reference: No co-morbidity			
Disability Onset 0-4 years ago	1.421*	1.409*	1.421**
	(1.103, 1.832)	(1.094, 1.816)	(1.102, 1.833)
Disability Onset 5-9 years ago	1.307**	1.292**	1.312*
	(1.065, 1.603)	(1.053, 1.585)	(1.068, 1.612)
Disability Onset10-15 years ago	1.190+	1.184	1.199
	(0.999, 1.430)	(0.984, 1.424)	(0.996, 1.443)
Reference: >15 years ago			
Income per CU < 500 Euro		1.199*	1.194*
		(1.008, 1.425)	(1.001, 1.424)
Income per CU 500-750 Euro		0.973	1.000
		(0.824, 1.148)	(0.839, 1.560)
Reference: >750 Euro			
Incomplete Educ.		1.294+	1.248
		(0.997, 1.680)	(0.960, 1.623)
Primary Educ.		1.051	1.023
		(0.794, 1.391)	(0.772, 1.356)
Reference: Secondary/Higher Educ.			
Widowed			1.103
			(0.929, 1.310)
Div./Single			0.888
			(0.668, 1.181)
Reference: Married			
No close relatives			1.546+
			(0.965, 2.476)
Close kin in same HH			1.180
			(0.893, 1.560)
Close kin in same municipality			1.165
			(0.903, 1.502)
Reference: Kin in other geogr. area			
> 2 people in HH			0.932
			(0.762, 1.140)
Reference: 1-2 people in HH			
Observations	2,058	2,058	2,058
Log Likelihood	− 3125.433	−3117.161	−3113.609
LR Test (df total/additional)	39.915*** (df = 7/7)	16.543** (df = 11/4)	7.104 (df = 17/6)
AIC	6264.865	6256.322	6261.218
Note:			+p<0.1; *p<0.05; **p<0.01; ***p<0.001

Table 3.2 Gompertz PH Regression Models - Males with Early Severe Disability Trajectory

	Dependent variable: Relative risk of dying		
	(1)	(2)	(3)
No daily activity	1.558***	1.584***	1.518***
	(1.247, 1.826)	(1.316, 1.905)	(1.258, 1.831)
Reference: Daily Active			
Suffers from multiple diseases	1.129	1.148	1.180
	(0.848, 1.339)	(0.918, 1.436)	(0.942, 1.477)
Reference: No Co-morbidity			
Disability Onset 0-3 years ago	1.590*	1.610*	1.603*
	(1.124, 2.248)	(1.139, 2.277)	(1.131, 2.272)
Disability Onset 4-9 years ago	1.640*	1.648**	1.610**
	(1.239, 2.171)	(1.245, 2.183)	(1.211, 2.140)
Disability Onset 10-14 years ago	1.315*	1.308	1.329+
	(1.002, 1.726)	(0.993, 1.713)	(0.992, 1.726)
Reference: >14 years ago			
Income per CU < 500 Euro		0.804	0.753*
		(0.617, 1.049)	(0.570, 0.995)
Income per CU 500-750 Euro		1.023	1.037
		(0.834, 1.255)	(0.842, 1.279)
Reference: >750 Euro			
Incomplete Educ.		0.971	0.972
		(0.753, 1.250)	(0.672, 1.256)
Primary Educ.		1.002	1.009
		(0.771, 1.302)	(0.774, 1.315)
Reference: Secondary/Higher Educ.			
Widowed			0.935
			(0.672, 1.303)
Div./Single			1.312
			(0.962, 1.788)
Reference: Married			
No Close Relatives			1.459
			(0.807, 2.637)
Close kin in same HH			1.289
			(0.886, 1.873)
Close kin in same municipality			1.191
			(0.880, 1.612)
Reference: Kin in other georgr. area			
> 2 people in HH			0.902
			(0.651, 1.251)
Reference: 1-2 people in HH			
Observations	857	857	857
Log Likelihood	−1671.067	−1669.136	−1664.587
LR Test (df= total/additional)	34.722*** (df = 7/7)	3.862 (df = 11/4)	9.097 (df = 17/6)
AIC	3356.134	3360.272	3363.175
Note:			+p<0.1; *p<0.05; **p<0.01; ***p<0.001

The estimates for the monthly income per CU show a between 12 and 15 percentage point increased hazard for female survival in the lowest income category. There is also a strongly significant effect of low levels of education. Women without completed education are found to experience a 11.4 to 14.5 percentage point higher risk of dying when compared to those with secondary or higher education. The results further suggest a higher risk of widowed and divorced or single women who are compared with married women (about 8% for widows and about 33% for singles and divorced women). The estimated hazards are 9.5 percentage points higher for widows and 36.7 for divorced and singles respectively. Furthermore, estimates suggest a 10.6 percentage points higher hazard for those who live in a household with more than one other person. Similar to the females, results show a between 15.8 and 16.6 percentage point higher relative risks for men who live in household with low, reported income per CU (less than 500 Euro per month) when compared to those in the highest income category. The estimates, shown in Table 3.6, further suggest that males without completed education are exposed to a between 17.3 and 19.4 percentage point higher risk than those with secondary or higher education. The impact of income and low education on male mortality when compared to females are relatively difficult to interpret as the associations with mortality are indirect, very complex, and has changed vastly over time. Widowed men are estimated to have a about 9 percentage point higher mortality risks while singles and divorced men appear to be exposed to a 42.5 percentage point higher hazard when compared to their married counterparts. The difference in the effect of widowhood between the sexes may be a relic of difference in age-specific mortality. As women survive on average longer than men in all subpopulations of this study, the are also exposed to a much higher risk to experience the loss of a partner and the negative effects of this event.

Table 3.3 Gompertz PH Regression Models - Females with Late Severe Disability Trajectory

	Dependent variable: Relative risk of dying		
	(1)	(2)	(3)
No daily activity	1.507***	1.508***	1.501***
	(1.353, 1.679)	(1.353, 1.682)	(1.346, 1.675)
Reference: Daily Active			
Suffers from multiple diseases	1.182**	1.183**	1.185**
	(1.059, 1.318)	(1.060, 1.319)	(1.062, 1.323)
Reference: No Co-morbidity			
Disability Onset 0-1 years ago	1.043	1.046	1.045
	(0.915, 1.189)	(0.917, 1.193)	(0.916, 1.192)
Disability Onset 2-3 years ago	0.930	0.932	0.933
	(0.804, 1.075)	(0.807, 1.078)	(0.806, 1.080)
Disability Onset 3-7 years ago	0.880	0.882	0.878
	(0.766, 1.009)	(0.768, 1.013)	(0.765, 1.008)
Reference: >7 years ago			
Income per CU < 500 Euro		0.971	0.973
		(0.866, 1.088)	(0.868, 1.092)
Income per CU 500-750 Euro		1.055	1.026
		(0.949, 1.173)	(0.903, 1.145)
Reference: >750 Euro			
Incomplete Educ.		0.849	0.849
		(0.712, 1.014)	(0.711, 1.016)
Primary Educ.		0.830+	0.831
		(0.688, 1.001)	(0.689, 1.003)
Reference: Secondary/Higher Educ.			
Widowed			0.903
			(0.802, 1.102)
Div./Single			1.069
			(0.876, 1.304)
Reference: Married			
No Close Relatives			1.021
			(0.785, 1.328)
Close kin in same HH			1.163
			(0.962, 1.406)
Close kin in same municipality			1.146
			(0.955, 1.374)
Reference: Kin in other geogr. area			
> 2 people in HH			0.949
			(0.845, 1.064)
Reference: 1-2 people in HH			
Observations	2,498	2,498	2,498
Log Likelihood	−5727.379	−5724.584	−5719.425
LR Test (df= total/additional)	73.172*** (df = 7/7)	5.589+ (df = 11/4)	10.319 (df = 17/6)
AIC	11468.76	11471.17	11472.85

Note: +p<0.1; *p<0.05; **p<0.01; ***p<0.001

Table 3.4 Gompertz PH Regression Models - Males with Late Severe Disability Trajectory

	Dependent variable: Relative risk of dying		
	(1)	(2)	(3)
No daily activity	1.524***	1.527***	1.502***
	(1.336, 1.739)	(1.337, 1.744)	(1.314, 1.715)
Reference: Daily Active			
Suffers from multiple diseases	1.165*	1.166*	1.171*
	(1.007, 1.348)	(1.007, 1.348)	(1.007, 1.350)
Reference: No Co-morbidity			
Disability Onset 0-1 years ago	1.215*	1.222*	1.245*
	(1.017, 1.451)	(1.022, 1.461)	(1.041, 1.490)
Disability Onset 2-3 years ago	1.130	1.135	1.140
	(0.925, 1.385)	(0.925, 1.391)	(0.929, 1.399)
Disability Onset 4-6 years ago	1.070	1.072	1.097
	(0.871, 1.314)	(0.872, 1.318)	(0.892, 1.350)
Reference:>6 years ago			
Income per CU $<$ 500 Euro		0.948	0.926
		(0.792, 1.136)	(0.769, 1.115)
Income per CU 500-750 Euro		1.020	1.058
		(0.881, 1.181)	(0.909, 1.231)
Reference: >750 Euro			
Incomplete Educ.		1.029	1.018
		(0.852, 1.243)	(0.841, 1.231)
Primary Educ.		1.060	1.053
		(0.865, 1.300)	(0.858, 1.292)
Reference: Secondary/Higher Educ.			
Widowed			0.981
			(0.836, 1.515)
Div./Single			1.228
			(0.919, 1.641)
Reference: Married			
No Close Relatives			1.531+
			(1.043, 2.247)
Close kin in same HH			1.256
			(0.978, 1.614)
Close kin in same municipality			1.207
			(0.973, 1.496)
Reference: Kin in other geogr. area			
$>$ 2 people in HH			0.957
			(0.799, 1.147)
Reference: 1-2 people in HH			
Observations	1,102	1,102	1,102
Log Likelihood	−2630.756	−2630.219	−2625.306
LR Test (df= total/additional)	49.022*** (df = 7/7)	1.072 (df = 11/4)	9.827 (df = 17/6)
AIC	5275.512	5282.439	5284.612

Note: +p<0.1; *p<0.05; **p<0.01; ***p<0.001

Table 3.5 Gompertz PH Regression Models - Females Free of Disability

	Dependent variable: Relative risk of dying	
	(1)	(2)
Income per CU $<$ 500 Euro	1.153***	1.122***
	(1.063, 1.251)	(1.034, 1.218)
Income per CU 500-750 Euro	1.005	1.029
	(0.932, 1.084)	(0.951, 1.113)
Reference: $>$ 750 Euro		
Incomplete Educ.	1.114***	1.145***
	(1.011, 1.226)	(1.038, 1.262)
Primary Educ.	0.955	0.981
	(0.865, 1.055)	(0.888, 1.085)
Reference: Secondary/Higher Educ.		
Widowed		1.078^{+}
		(0.998, 1.165)
Div./Single		1.329***
		(1.196, 1.478)
Reference: Married		
$>$ 2 people in HH		1.106***
		(1.031, 1.187)
Reference: 1-2 people in HH		
Observations	30,121	30,121
Log Likelihood	$-$17851.13	$-$17834.87
LR Test (df= total/additional)	3934*** (df = 6/6)	32.5231*** (df = 9/3)
AIC	35714.27	35687.74
Note:	$^{+}$p$<$0.1; *p$<$0.05; **p$<$0.01; ***p$<$0.001	

Table 3.6 Gompertz PH Regression Models - Males Free of Disability

	Dependent variable: Relative risk of dying	
	(1)	(2)
Income per CU $<$ 500 Euro	1.156***	1.168***
	(1.074, 1.245)	(1.082, 1.261)
Income per CU 500-750 Euro	1.027	1.073***
	(0.972, 1.086)	(1.013, 1.136)
Reference: >750 Euro		
Incomplete Educ.	1.193***	1.174***
	(1.113, 1.278)	(1.095, 1.258)
Primary Educ.	1.080***	1.071*
	(1.010, 1.156)	(1.001, 1.146)
Reference: Secondary/Higher Educ.		
Widowed		1.096***
		(1.009, 1.191)
Div./Single		1.425***
		(1.319, 1.541)
Reference: Married		
$>$ 2 people in HH		0.982
		(0.928, 1.038)
Reference: 1-2 people in HH		
Observations	29,157	29,157
Log Likelihood	−27459.53	−27420.23
LR Test (df= total/additional)	5889.11*** (df = 6/6)	78.594*** (df = 9/3)
AIC	54931.05	54858.46
Note:	$^{+}$p<0.1; *p<0.05; **p<0.01; ***p<0.001	

3.4 Discussion

The onset of old-age disability is a critical life course transition associated with the loss of autonomy and quality of life [18, 45, 26]. Not only since the influential work by Fries and colleagues it has been closely linked to death [157]. In public health research mortality measures such as life expectancy are not rarely used as proxy for population health [cf. 180, 186]. As social, technical, and medical progress have led to vast improvements in the lives of people with limitations, the association between the onset of disability and death might have changed. Better screening, prevention, and treatment of formerly limiting diseases and chronical conditions have increased the chances, especially of mildly disabled people, to participate in everyday social activities and remain relatively independent [271]. On the other hand, it can be argued that such a buffering effect of increased functionality in mildly disabled individuals can hardly mitigate the immense increase in prevalence of disability due to ongoing population aging. While medical and social progress have also led to higher chances of surviving formerly deadly diseases and conditions, this might have unlocked a hidden burden of disability in older ages [20, 94].

This development also has implications on how disability is assessed and related to mortality. Independently of expansion or compression of morbidity, ongoing population aging will lead to a growing proportion of individuals with disability [56]. To improve future health equity, prepare public health interventions, and, better predict future health care costs, it will advantageous to know if there are common pathways after the onset of disability, how they are related to mortality, and how this relationship changes over time.

In an attempt to provide a replicable methodology for identifying such shared disability trajectories, which then can be used for further analyses, this work describes how state sequences based on the numbers of limitations and onset times of the first three disability events can be used to feed an adapted optimal matching algorithm. This algorithm compares individual sequenced disability experiences by onset age and length of state durations to identify shared pathways. Assessment tools like the ASW allow to reduce the number of possible outcomes with some degree of confidence. The analysis of mortality disparities within identified pathway groups in this work shows how these techniques can help to reduce complexity for example in the assessment of health inequalities after onset of disability.

Three pathways are identified for both sexes that differ by severity and timing of first onset. In contrast to a recently published work on disability pathways in England, there was a group with consistently low levels of functional limitations but no medium impairment

group [232]. Individuals with disability who were not selected in the *mild* pathway, mostly experienced the onset of severe limitations at one moment or within a short time frame and are only distinguished by the average onset age at first event.

The analysis of socioeconomic factors contributing to survival differences within groups adds to the better understanding of the evolving relationship between disability and death [121, 279]. Although limited to a relatively small set of variables, the estimated mortality hazards suggest that survival is determined by different impact factors in different pathway groups. While there was no evidence of a socioeconomic gradient in any of the three groups, the results suggest small differences in the importance of performance and medical measures.

The most robust finding is the positive effect of daily activity, which is found to decrease mortality risks between 30 and 50 percentage points. Inactive men in the *early severe* pathway group were exposed to a higher risk compared to females. The differences between the sexes in the late severe group are found to be small but larger than in the other groups. The largest relative risks can however be found for inactive males in the *mild* group who are estimated to have an about 63 percentage point higher risk than those who are active. Although sedentary lifestyle and the lack of daily activity have been repeatedly linked to early onset of disability and even mortality, the importance of this finding cannot be understated [181, 219, 17]. Among others, these results suggest an underlying health promoting effect of community and an active social network. While the ability to perform daily activities is a condition for maintaining friendships and other personal relationships, it has frequently be shown that an active social network gives incentives to be active on a daily basis and improves the health of older adults [cf. 75, 95].

Higher relative mortality risks in the *late severe* group suggest increasing opportunity costs of adaptation with older ages. While the *early severe* group might have adapted and found alternative activities, it is presumably more difficult to adapt when limitations occur later in life. As the decline of physical activity generally accelerates with older ages, it will be important to identify these high-risk groups and find ways to postpone or slow down this development [43]. While a more accessible build-in environment was found to peak daily activity in older adults and would be desirable, small incentives such as diaries and digital step counter can also efficiently boost the levels of daily activity often even independent of functional limitations [268, 196, 188].

As mentioned above, the effects of time spent with disability on mortality are difficult to interpret. The results generally support the adaptation hypothesis, especially within the *early severe* group. For both sexes, there appears to be a gradient with the highest mortality risk in the years right after onset while those who have spent the longest time in disability are exposed to the lowest risk. Although more complete longitudinal information is needed to exclude the possibility that the decreasing risk is a pure relic of selection upon survival, these results indicate that it might be important target those with an early onset of severe old-age disability and help them to better adapt to a live with disability.

Once individuals experienced the onset of disability, their social position appeared to have only minimal impact on their survival, while well-known mitigating effects of high income and higher education are prevalent in the disability-free population. Even though the disability-free population is on average younger and medical predispositions remain unobserved, this finding might be highly relevant for the research on health inequalities. The comparison of age-specific survival probabilities suggest that socioeconomic effects play a larger role for the occurrence of old-age disability than on mortality after the onset of a chronic condition. Disability appears to have an equalizing effect while the occurrence of limitations is highly dependent on socioeconomic factors [cf. 278]. To assure that these effects are not caused by health selection, more research is needed that further disentangles the influence of socioeconomic context on disability and survival.

Despite the use of a newly linked data source, there are several limitations due to the structure of the data and methodological choices. As there is just one cross-sectional time point from which information on disability and socioeconomic measures is obtained, it cannot be ruled out that these levels have changed substantially during the follow-up period. Especially individuals who were assigned to the disability free or *mild* category might have experienced further events that are not captured but would have placed them in one of the other categories. Thus, as many individuals in the disability-free subpopulation might have experienced an unobserved onset of disability after the survey year, mortality rates of those who never suffered a disability according to the definition may have even be lower. On the other hand, individuals living in a household with a dependent or disabled relative are often found to be exposed to higher levels of stress which negatively affect their health and ultimately their risk of dying [294, 250]. Therefore, the disability-free population may not resemble a total population sample and mortality differences might be even larger.

Moreover, there are numerous algorithms and methods that allow for grouping age-sequenced trajectories. Some of these methods do not produce reliable results or place the same individuals in different trajectory groups when the process is repeated [305]. Algorithms for the comparing state sequences are always based on a simplification of the data that might influence the assignation of individuals to one group or another. Therefore, selecting the right method might not always be straightforward and depends on the focus of the analysis. As sequences may be more similar regarding one aspect, such as duration, they may differ more when focusing on another [274]. With the focus on dissimilarities by the time spent in a distinct successive disability state, it can for example not be excluded that small differences in the selected population would have led to different assignation of individuals to the *early* or *late severe* group, especially where the onset age was close to the threshold age where these groups are separated.

The validity of the results might further be affected by the aforementioned selection mechanism. It is possible that selection upon survival has biased the risk to be part of the study population and the different pathway groups. It can further be assumed that when information on dates is gathered in face-to-face interviews with either very old people or representatives, there will be a certain recall bias and possibly some digit preference with respect to onset dates [cf. 266]. As the age at onset of disability was derived from this information it might be affected by these biases.

Nevertheless, the results confirm that there are underlying shared mechanisms that allow for grouping disability trajectories that then can be used to analyze the implications of a certain disability history on various outcomes. The comparison of within group survival differences further confirms the applicability and results suggest that socioeconomic differences appear to be minimal after onset of disability. Furthermore, it could be shown that the importance of daily activity, co-morbidity, and time after onset differs by disability trajectory.

This research needs to be expanded on to answer questions like what determines the ability to adapt to certain disabilities and how does mental health relate to these processes. The assessment of different trajectories across different health states or markers for different pathways through disability will remain important. From an economical perspective, it might help to better predict future care costs. If, for example, the distribution of individuals in different pathways would change from time a to time b, it will affect average health spans. If there were for example relatively more individuals with a *late severe* pattern, this would change many factors including the demand of care just due to their higher mortality compared

to the other pathway groups. At an individual level further assessment of this kind might help to make personal decisions regarding care or more general about how to manage different domains of life after onset of disability [99]. A better understanding of what to expect in certain (and in this case dire) situations has the potential to contribute to a better quality of life.

Chapter 4

Urban environment and mortality disparities in Andalusia

Abstract

Background: Analyses of health and mortality disparities between today's urban and rural populations appear to be exclusively focused on vastly urbanising countries. By incorporating environmental data at census tract level and accounting for within-area homogeneity, this work attempts to extend classic rural–urban comparisons.

Method: Geographical information is linked to a register-based mortality follow up and Spanish census data for the autonomous community of Andalusia. We then apply mixed effects Cox proportional hazards models to estimate individual mortality differences and account for area variations between residential areas.

Results: Estimated effects suggest that the shared degree of "urbanicity" does not affect individual hazards of mortality, whereas environmental- and population-based measures influence the relative risk of dying despite controlling for individual-level risk factors.

Conclusion: Although we do not find an impact of physical urban measures, our results reveal persistent that area-related mortality disparities which can help to explain the mechanisms behind prevalent spatial-temporal inequalities such as those in Andalusia.

4.1 Introduction

According to the latest United Nations report on urbanization, the proportion of the world population living in urban areas is expected to increase from 55% in 2018 to 68% by the year 2050 [285]. Accelerated growth of cities predominantly in developing countries will trigger various changes in social structure, occupational activities, and distribution of wealth. Such a development also entails great challenges regarding social equity and long-term development of public health [8, 312].

Recent public health research suggests that various unfavorable health conditions including obesity, high cholesterol, and different forms of mental illnesses are more prevalent among individuals living in cities, when compared to their rural counterparts [147, 85, 289]. The relationship between health, mortality, and environmental features of residential areas, however, is complex. For example, exposure to environmental hazards might have a time-lagged effect on a person's health, yet residential areas are also constantly changing and developing within different cultural and social frameworks [253]. In the context of rural-urban differences, environmental effects could also work in both directions, as living in cities is associated with both harmful and health-promoting features at the same time. For example, the increasing adaption of unfavorable diets in China's urban population raises concerns about future health disadvantages. However, vaccination rates, hygiene, and access to health care are significantly better in China's urban centers than in most rural areas [116].

Traditionally, rural and urban subpopulations are compared by using aggregated data sources and dichotomous measures to distinguish between the rural and urban area types. This, however, fails to capture existing heterogeneity and interactions between environmental area features and the variable of interest, which are often masked by differences in the population composition [100]. Considering such hidden clustering effects and accounting for positive health aspects of urban life like the closer proximity to hospitals and specially qualified doctors, it may appear as if urban residents are more likely to have a health advantage over their rural counterparts.

However, urban dwellers or those who spend most of their day in a highly urbanized area are exposed to specific environmental risk factors like high levels of air pollution and the lack of access to green space [122, 225]. Numerous studies also link unfavorable health effects with increasing spatial proximity to areas with high exposure to negative environmental features, like natural gas wells or high relative amounts of Nano material [235, 60, 92].

Thus, health and well-being of individuals nested in neighborhoods or other forms of small areas are influenced by shared exposure to certain environmental stress factors. Such effects may occur independently of their individual characteristics and may be mediated through the degree of "urbanicity", referring to the degree or extent of urban features within residential areas [302, 315]. Even if the majority of health and survival disparities can be traced to behavioral, socioeconomic, or biological differences, accounting for exposure to environmental hazards complements the analysis of individual risk factors [303].

To our knowledge, only a handful of studies have addressed disparities between rural and urban populations in large regions of high-income countries, possibly due to much lower growth rates when compared to cities in China and other vastly urbanizing countries. Even if these changes are modest, attractive job markets in or near urban centers and the growing demand for service work have led to a continuously changing population composition between rural and urban settings with regard to age distribution, education, and other wealth parameters. Consequently, population movements and area developments lead to environmental changes which in turn affect the population's health [202, 120, 245]. Although comprehensive health insurance coverage in most high income countries prevents large-scale health and survival disparities, the examination of indirect environmental risk factors, some of which may be specifically urban or rural can help to explain and ideally prevent the evolving health and mortality gap between social groups. Particularly in places like Southern Europe, it is necessary to analyze phenomena like urban heat islands where the asphalt and other artificial surfaces store and accumulate summer heat during the day and thereby create a substantially warmer environment during the night for individuals residing and working in urban centers [211].

In the context of this work, we aim to contribute to the debate on urban-rural differences particularly in the field of small area analysis by estimating the effects of environmental impacts and urban characteristics on individual-level survival over time. First, we introduce the conceptual framework focusing on the measurement of "urbanicity." Second, we explain in more detail the construction of our index to measure the latent concept "urbanicity" and introduce the data infrastructure. Third, we fit a Cox proportional hazards (PH) model with mixed effects to estimate individual-level mortality risks and the effects of exposure to urban environment over time. We account for possible stratum homogeneity by including random area effects. Finally, we compare our results to alternative models and discuss the main findings and limitations.

4.1.1 Measuring "urbanicity" and small area environment features

The increasing availability of spatially referenced data, greatly improved computing power, and the development of more advanced statistical approaches help to generate new strategies for comparing rural and urban subpopulations. Classic studies often apply dichotomous indicators to distinguish between urban and rural areas. They commonly rely on a core set of characteristics aggregated at differently specified areas, most commonly based on administrative boundaries [302].

A dichotomous classification according to administrative boundaries appears to be straightforward, but fails to capture part of the between-area variation and relevant urban characteristics related to infrastructure, geographic position, and distribution of active space [68]. In fact, cities are often surrounded by heavily populated areas which might not be part of the same administrative unit but are still exposed to similar conditions.

To address calls for a more nuanced approach to the subject matter [193], Vlahov and Galea [301] proposed a refined conceptualization for measuring urban space. The crucial step of their approach was to disentangle the two related and often synonymously used concepts urbanization and "urbanicity." While they defined urbanization as a process of growth of cities in terms of area and population over time, they related the term "urbanicity" to a current state of an area which can differ by degrees of certain urban characteristics like the proportion of built in environment. In other words, their concept captures the "nature of urban environments" [70]. "Urbanicity" is a concept strongly dependent on the regional context and changes over time [169]. As such, it is difficult to define. The lack of a universal definition, however, provides the opportunity to propose new measures that can then be compared over time.

4.2 Materials and methods

4.2.1 Data

The Base de Datos Longitudinal de Población de Andalucía (BDLPA) is a comprehensive, longitudinal data infrastructure containing administratively collected information on individuals who were registered in the autonomous community of Andalusia at the time of the Spanish population and housing census of 2001. The around 8.3 million individuals are biannually followed up on deaths and emigration based on data derived from regional population registers. A 10% sample of BDLPA data spells can be accessed through the website of the

Institute of Statistics and Cartography of Andalusia (IECA) . The advantage of the BDLPA is that it can be linked to other administrative data, thus allowing us to geo-reference all subjects from the 10% sample and group them into census tracts. To ensure anonymity of individuals and single households, the information was not accessible below census tract level according to the definition of census of 2001. Contextual geographical information is obtained from the CORINE land-cover raster database (Coordination of information on the environment) while the cartography unit of the IECA provided maps for the geographical analysis.

4.2.2 Indicator of urban environment

The degree of "urbanicity," or in other words the degree of how urban an area is, refers to a multidimensional and latent concept. Accounting for the complexity of such a concept, we choose a mix between a theoretical and data-driven approach to construct a multicomponent index that allows us to distinguish between different degrees of urban environments [cf. 147, 241, 70, 302]. By examining graphical tests and correlation coefficients between all accessible environmental variables related to urban settings, we identify four main contributors to the latent concept "urbanicity", depicted in Figure 1 below.

First, we calculate the population density per census tract as a measure of relative crowdedness and standardized observations based on the overall deciles to make them comparable with other scale components. Second, we calculate and add the average coverage with medical service based on estimated service area polygons that represent the distance that can be covered between a health facility and any point on the map within 30 minutes driving-to-facility time. Third, we use satellite imagery data from the CORINE land cover database to estimate the artificial surface area per census tract for the year 2006. Fourth, we obtain and add road density by estimating the total length of line objects (transportation networks) within area units (km of road per sq.km of land surface). The weights with which single components enter the index are estimated through maximum likelihood factor analysis incorporating standardized single component values [cf. 148, 41]. Factor weights are represented in Appendix Table 1. The resulting index variable is further normalized and centered around zero. A Crohnbachs alpha score of 91% suggests sufficient internal consistency of all indicator components.

A graphical quality test of the index is depicted in Figure 4.2 which displays a choropleth map of Seville, the biggest city, population-wise, in Andalusia [141]. Scores for the multicomponent indicator for urban environment are represented by census tract and darker

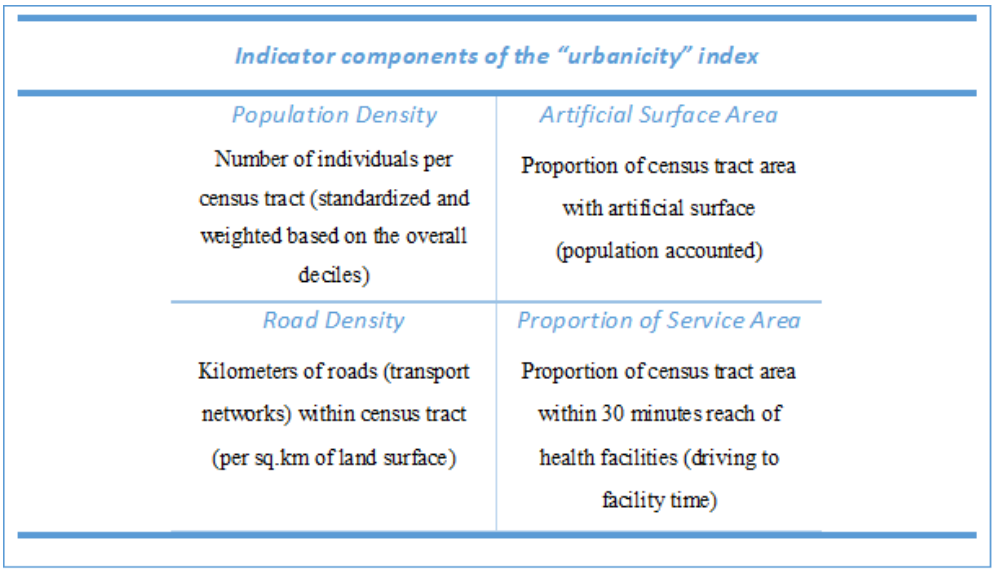


Fig. 4.1 Components of the Urbanicity Indicator

shades are associated with a higher degree of "urbanicity." Areas identified as urban are rather small and located in the city center, which indicates a good graphical fit for our measure.

4.2.3 Environmental and population-based area features

The rather physical "urbanicity" indicator does not allow us to capture area-specific heterogeneity regarding environmental hazards and potentially harmful population or social features unrelated to "urbanicity." As we attempt to distinguish between different kinds of urban and rural areas, it is important to account for further heterogeneity between areas similar to what we can observe in reality. To capture between-area differences regarding the latent socioeconomic situation and the exposure to area-specific environmental hazards, we incorporate additional population measures and aggregated survey answers on the residential environment in the analysis. After performing sensitivity tests, we choose perceived average cleanliness, noise exposure, and air pollution to represent additional environmental hazards in our models. The proportion of unemployed individuals at working age and the proportion of single households enter the models as population features. All variables are standardized with reference to the mean values for Andalusia in the year 2001.

4.2.4 Study population and individual-level information

There are two main sources for individual-level information. The BDLPA is a quasi-continuous mortality follow-up, which is semiannually updated and corrected. By using an

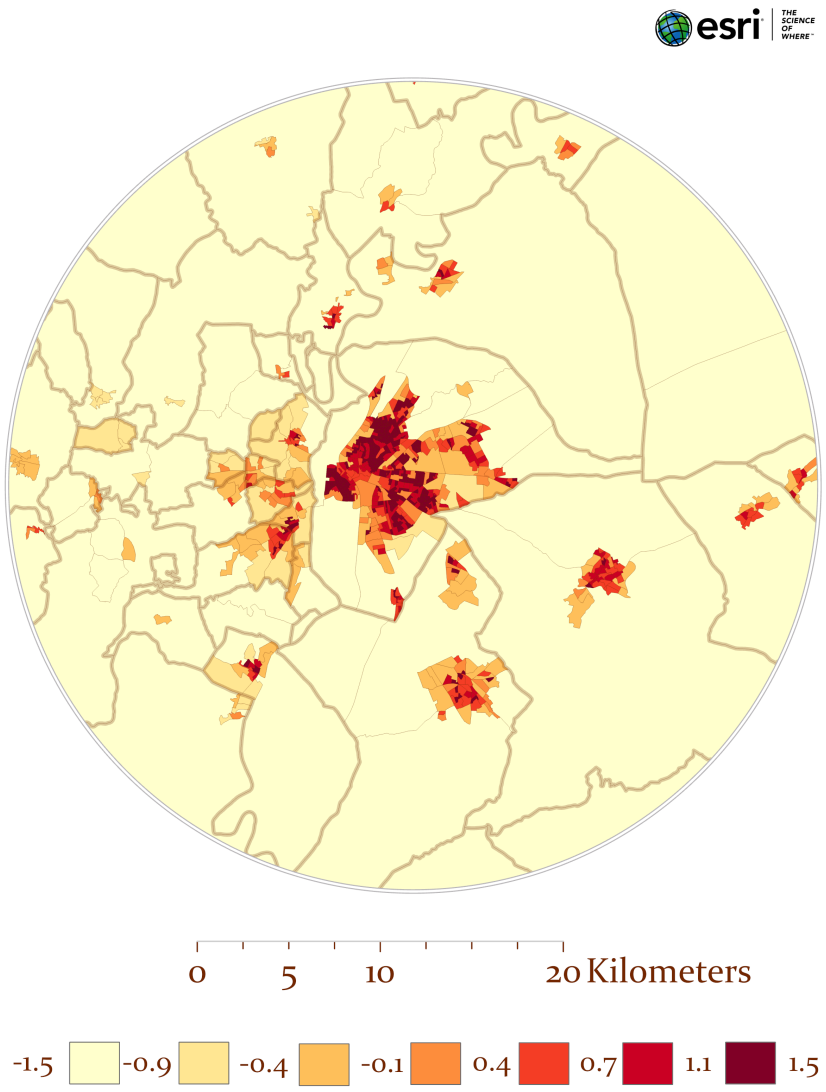


Fig. 4.2 City of Seville - Scores for the Urban Environment Indicator

individual identifier, we are able to link all subjects in the study to their answers from the population and housing census in 2001, the baseline year for our analysis, and follow them up until 2014. Individuals who died or emigrated within the follow up period are censored at the date of death or emigration. All others are right-censored at the end of the observation period. Information on the residential area and the individual socioeconomic context are only available for the baseline year of the study (2001). To reduce the bias induced by potentially unobserved changes in residence and other individual time-varying information, only individuals between the ages 35 and 80 are included in the analysis. The selection of the age groups is premised on informed assumptions about living conditions and tenure status derived from commonly observed life course trajectories in the given age range in Spain [230]. In general, individuals between age 35 and 80 have resided in their house or apartment for relatively long periods and are, as they grow older, more likely to own their dwelling. The probability of moving is rather low for those age groups and individuals are more likely to be exposed to the same environment for the time of our study.

Limiting the age range was further motivated by the distribution of the event of interest as more than 90% of all deaths in Andalusia occur after age 35 but before age 92, the highest age individuals reach at the end of the follow up period in 2014 [cf. 207, 298]. A consequence of selecting these age groups is that the sample size decreases from 723,234 to 351,769 individuals. We tested all models with different age ranges but did not observe greatly different patterns in population composition or general model outcomes. To assure that neither the observed population nor the area-specific "urbanicity" and environmental characteristics have dramatically changed over 13 years of observation, particularly in the light of substantial economic fluctuations since 2001 [91, 88], we perform sensitivity tests for different time periods and with different age ranges.

One strength of this analysis lies in the combination of individual-level information with area-specific factors captured in a multilevel setting. Such data structure guarantees that possible index effects are not caused by unobserved heterogeneity between subpopulations. We included sex, disability status, and marital status observed at the baseline year as individual-level variables in our models. To control for socioeconomic individual differences, we incorporated several indicators representing social position as for example the highest educational degree, ownership status of the dwelling, and car ownership status. All socioeconomic variables are derived from the census questionnaire of 2001.

4.3 Statistical approach

The incorporation of area effects into an analysis of individual-level mortality differences requires statistical testing for potential impacts of cluster-specific effects, and in case of geographical data the spatial distribution. The graphical representation of the multicomponent "urbanicity" index and statistical tests regarding the environmental variables suggest that observations are more likely to be similar if they are geographically closer to each other. To assess if observations are spatially autocorrelated, intrinsic stationarity is assumed before we calculate a row-standardized matrix of spatial weights that is based on the list of contiguous neighbors. At least one point of the boundary of a spatial polygon that represents a census tract has to be within snap distance of at least one point of a neighboring polygons' border to meet our contiguity condition [cf. 13, 32]. We then calculate the product-moment correlation coefficient (Moran's I) as statistical test for spatial autocorrelation [264, 201]. While spatial autocorrelation of mortality indicators would justify the analysis of area differences in the first place, we also estimate Moran's I for other central variables including the index for "urbanicity" and aforementioned area-specific environmental features.

4.3.1 Statistical model

Following the descriptive tests for spatial autocorrelation, we estimate mortality disparities by degree of "urbanicity" and environmental impact with an extended version of the Cox PH model. The original model is the most commonly used approach to model censored time to event data, particularly when the main interest is to obtain relative effects of covariates [197]. Such effects are assumed to be proportional over time and enter the model multiplicatively as expressed in the following equation [cf. 156].

$$h(t) = h_0(t) \exp(\beta_i X_i) \tag{4.1}$$

, where h_0 is the baseline hazard and $\exp(\beta_i X_i)$ the non-negative function of covariates. Hazard ratios are obtained through the maximization of the partial log likelihood with respect to $\beta_i i X_i$ [10, 283]. Since only the right-hand side of the formula is maximized, the Cox PH model does not require you to specify the underlying baseline distribution. Due to our assumption that individuals are nested in small areas where they are exposed to similar environmental hazards and the same degree of "urbanicity," we choose to apply an extended version of the Cox PH model, which allows us to account for such homogeneity within clusters. Thus, a stratum-specific frailty term is added to the original model [cf. 16]. The resulting Cox PH model with mixed effects can be considered as shared frailty model with a

normally distributed stratum specific frailty term Z_j as follows [281].

$$h(t) = h_0(t) \exp(X_i\beta + Z_j) \tag{4.2}$$

, where Z_j is the design matrix for random effects which captures homogeneity within clusters. The model can be interpreted as multilevel survival model with shared frailties. The added random effect term can be understood as relative effect of given covariate patterns on the baseline hazard which varies across census tracts [223]. Given the set-up of our analysis, it is necessary to account for left truncation [44]. This adjustment affects survival estimates for everybody in the sample because their time under risk of dying before the start date of the study remains unobserved. In other words, we select individuals based on their survival upon the start year of the examination. To account for left-truncation and assure we measure age-specific mortality differences, we choose to use person years as the time scale in our models. Calendar time and cohort effects are accounted for by including birth cohort effects as covariate [49].

4.4 Results

In order to analyze area differences, spatial variation must be present in the variable of interest. We determine to what extent these differences are spatially associated by estimating Moran's I for the variables in our analysis, which can be interpreted as the correlation between a variable and its spatial neighbors [14]. These estimates and associated p-values are presented in Table 1. All variables of interest exhibit significant spatial autocorrelation, justifying our analysis of area differences in Andalusia.

As the analysis aims to highlight the impact of shared environmental factors on individual-level survival, the population under observation is considered to be nested in geographical units, which requires the application of a multilevel model structure. The estimated coefficients (fixed effects) of four separate Cox PH models with mixed effects and a step-wise increasing number of covariates are presented in Table 4.2 and 3. When compared using Likelihood Ratio Tests, these four presented models fit significantly better than a Cox PH model without random effects. Furthermore, as these models include more explanatory variables, fit improves.

The estimated hazard ratio in Model 1 indicates that alone, the selected urban features do not substantially affect individual mortality differences between census tracts in Andalusia.

Table 4.1 Morans I Statistics for univariate relationships of indicators with geographic dimension

	Morans I Statistic	*p-value*
Smoothed SMR	0.380	< 0.001
"Urbanicity" Index	0.784	< 0.001
Perceived Noise	0.570	< 0.001
Perceived Neighborhood Pollution	0.444	< 0.001
% Single Household	0.568	< 0.001
% Unemployed	0.500	< 0.001

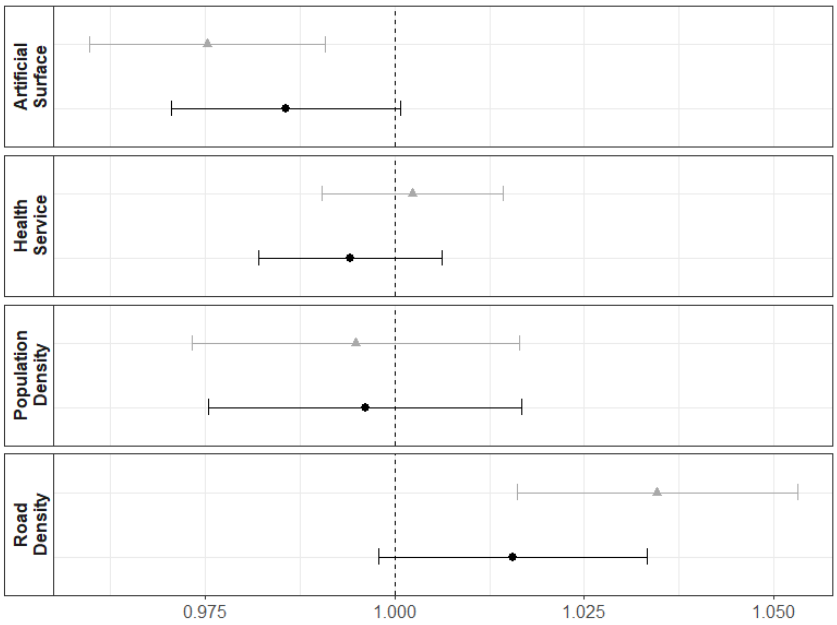


Fig. 4.3 Estimated Hazard Ratios (95% CI) for “Urbanicity” Indicator Components (grey - Model 1, black Model 4

In the second model, we account for within-area population heterogeneity by incorporating socioeconomic and demographic individual-level variables with well-documented indirect effects on mortality. The estimates for individual-level variables, depicted in Table 3, show typical mortality patterns. Men between ages 35 and 92 have a substantially higher relative mortality risk compared to their female counterparts. Estimates further suggest that individuals with functional limitations and other disabilities have a mortality hazard more than three times higher than those without these impediments. Moreover, all socioeconomic variables point towards increased relative risks for less wealthy and less educated individuals, with reference to both the more affluent and those with university education. As the estimated hazards for these individual-level risk factors change only marginally with the incorporation of environmental and population features, we present them in a separate table to avoid distraction from the effects of primary interest.

After including individual-level differences (Model 2), the effect of urban environment appears more pronounced than in Model 1. In Model 2, every unit increase in the "urbanicity" scale of a census tract increases the estimated hazard of dying by three percentage points. While changes in the individual-level impact factors are negligible between different models, the effect of the degree of "urbanicity" on survival varies with the incorporation of additional area-level characteristics. In Model 3, we incorporate the effects of perceived cleanliness, noise, air, and water pollution in an attempt to control for different kinds of heterogeneity between urban areas with the same degree of "urbanicity." The estimates suggest that including such environmental area features reduces the effect of the degree of "urbanicity" to 1.7 percentage points for every unit increase. Both cleanliness and pollution are found to have a highly significant but small effect on survival. Perceived cleanliness of the area is estimated to increase the hazard by 0.07 % for every unit of increase. Results also suggest the hazard increases by 0.2 % per every one additional percentage point of perceived pollution. In Model 4, when we include population characteristics of small areas, the estimated hazard for the "urbanicity" indicator is very close to one. As this "urbanicity" effect shrinks, we find that the proportions of both unemployed individuals and single households in a census tract increase individual hazards by 0.25 and 0.94 percentage points respectively. The addition of population-based measures does not appear to influence the effect of environmental area features.

We also estimate possible effects of single indicator components on the mortality hazards in a given census tract. These results (Figure 3) indicate that road density negatively affects survival in a model with mixed effects and without additional area variables, while

Table 4.2 Estimated Hazard Ratios from a Cox PH Model with mixed effects - Second level fixed environmental effects on individual hazards

	Dependent variable:			
	Hazard Ratios (95% CI)			
	(1)	(2)	(3)	(4)
"Urbanicity" Index	1.0061	1.0308***	1.0176**	1.0027
	(0.9963, 1.0158)	(1.0212, 1.0404)	(1.0058, 1.0295)	(0.9904, 1.0150)
Perceived Cleaness			1.0007**	1.0005*
			(1.0001, 1.0012)	(0.9999, 1.0011)
Perceived Pollution			1.0020***	1.0020***
			(1.0011, 1.0029)	(1.0012, 1.0029)
Perceived Noise			0.9999	0.9995
			(0.9991, 1.0008)	(0.9987, 1.0004)
% Unemployed				1.0025***
				(1.0013, 1.0036)
% Single HH				1.0094***
				(1.0070, 1.0119)
Individual Variables		x	x	x
Observations	351,769	351,769	351,769	351,769
Integr. Log Likelihood	−542,763.5	−538,537.7	−538,510.2	−538,471.8
LR Test	128.61 (df = 2)***	8451.6 (df = 12)***	55 (df = 15)**	76.8 (df = 17)***
Note:				*p<0.1; **p<0.05; ***p<0.01

the percentage of artificial surface has a slightly positive effect on survival. Just as with the index variable, the effect of single components vanishes when accounting for further environmental and social variables. Given that areas with the same level of "urbanicity" are still quite heterogeneous, the presumed negative effects of road density, urban contamination, and other explicitly urban risk factors appear to be less important in explaining mortality disparities between census tracts.

An advantage of shared frailty models over classic survival approaches is the ability to estimate relative effects of covariate patterns on the baseline hazard across clusters. Assuming subjects are exposed to shared environmental risk factors that, in spite of individual-level differences, influence their survival risk, we incorporate a normally distributed random effect for the residential area in the model. Estimated median frailties and its variations for all models are depicted in Figure 4. Naturally, the random variation is lower in models where we account for the shared additional area effects. Nevertheless, if translated into risk scores, there are still substantial differences between census tracts. For example, in Model 4, a census tract about one standard deviation above the mean corresponds to a relative mortality risk of 1.119. In other words, there is almost a 12 percentage point increase in the hazard of dying compared to the mean census tracts. Further variation measures can be found in Appendix Table 2, which we also provide as a summary of the likelihood ratio tests between all models and their counterparts without the additional random effects. The test statistics indicate that models with these shared frailties improve the fit significantly.

Figure 5 shows the effect of shared frailty by census tract. Values are presented in exponentiated form and can therefore be understood as risk scores. The darker the shade, the higher the unexplained relative mortality risk of individuals in the respective area. While the majority of census tracts experience mortality risks close to the mean (risk scores between 0.97 and 1.03), some hotspots appear to exist in central Andalusia, where mortality risk is up to 15 percentage points higher than average. Contrary to previous analyses of spatio-temporal differences, these high mortality areas do not appear to be clustered in the southwest area of the region but are instead spread randomly throughout Andalusia [207].

Table 4.3 Cox PH Model with mixed effects - Individual level variables corresponding to models in Table 4.2

	Dependent variable:		
	Hazard Ratios (95% CI)		
	(2)	(3)	(4)
Male	2.0872***	2.0883***	2.0898***
	(2.0682, 2.106)	(2.0692, 2.1074)	(2.0707, 2.1089)
Reference: Female			
Physically Dependent	3.0286***	3.0180***	3.0028***
	(2.9617, 3.0954)	(2.9512, 3.0848)	(2.9360, 3.0695)
Reference: No Dependency			
Single	1.412***	1.4126***	1.4038***
	(1.3826, 1.4422)	(1.3829, 1.4424)	(1.3740, 1.4336)
Widowed	1.1836***	1.1823***	1.1808***
	(1.1586, 1.2085)	(1.1573, 1.2073)	(1.1559, 1.2058)
Divorced/Separated	1.4794***	1.4765***	1.4768***
	(1.4249, 1.5340)	(1.4219, 1.5310)	(1.4222, 1.5314)
Reference: Married			
No or Incomplete Educ.	1.3798***	1.3837***	1.3942***
	(1.3387, 1.4210)	(1.3425, 1.4249)	(1.3529, 1.4357)
Primary/Secondary Educ.	1.1595***	1.1598***	1.1652***
	(1.1155, 1.2036)	(1.1157, 1.2038)	(1.1211, 1.2093)
Reference: Tertiary Educ.			
Does not Own House	1.1625***	1.1537***	1.1399***
	(1.1362, 1.1888)	(1.1274, 1.1801)	(1.1134, 1.1664)
Reference: Does Own House/Apartment			
Does not Own a Car	1.2690***	1.2693***	1.2653***
	(1.2493, 1.2887)	(1.2497, 1.2890)	(1.2456, 1.2851)
Reference: Does Own Car(s)			
Observations	351,769	351,769	351,769
Note:			*p<0.1; **p<0.05; ***p<0.01

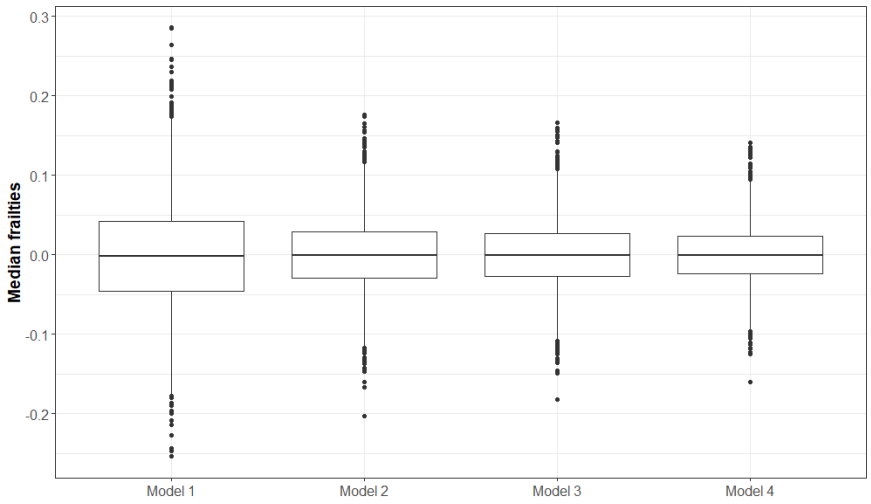


Fig. 4.4 Random area effects by model (median frailties)

4.5 Discussion

In this work, we aim to extend classic approaches for analyzing mortality disparities between rural and urban subpopulations. To capture the rural-urban gradient and estimate "urbanicity" effects on survival, we disentangle population and physical features from environmental impact factors in residential areas. We use satellite imagery data and census information to identify four universal predictors of "urbanicity," the multidimensional latent concept that describes the "urban nature" of an area [70]. Our index represents an improvement over classic binary measures that are based solely on population density. The use of census tracts as a clustering unit increases the precision with regard to area size and reduces the risk of misclassifying large areas as urban if, for example, only a part of the overall area exhibits urban features. Therefore, our approach offers an advantage over comparable measures, such as the "rurality index" proposed by Ocaña Riola and Sánchez-Cantalejo [208] in which data was aggregated at the municipality level.

We incorporate our index in mixed effects Cox PH models to estimate long-term survival according to different degrees of "urbanicity" along with individual and shared environmental risk factors. Results suggest that individuals residing in areas with higher levels of unemployment, single households, and perceived pollution face small but highly significant survival disadvantages, even after controlling for individual-level risk factors. While population-based and environmental factors appear to explain the majority of geographical survival differences in modern-day Andalusia, we found no clear evidence that physical urban environment, as captured through the aforementioned "urbanicity," index had an effect on survival. The

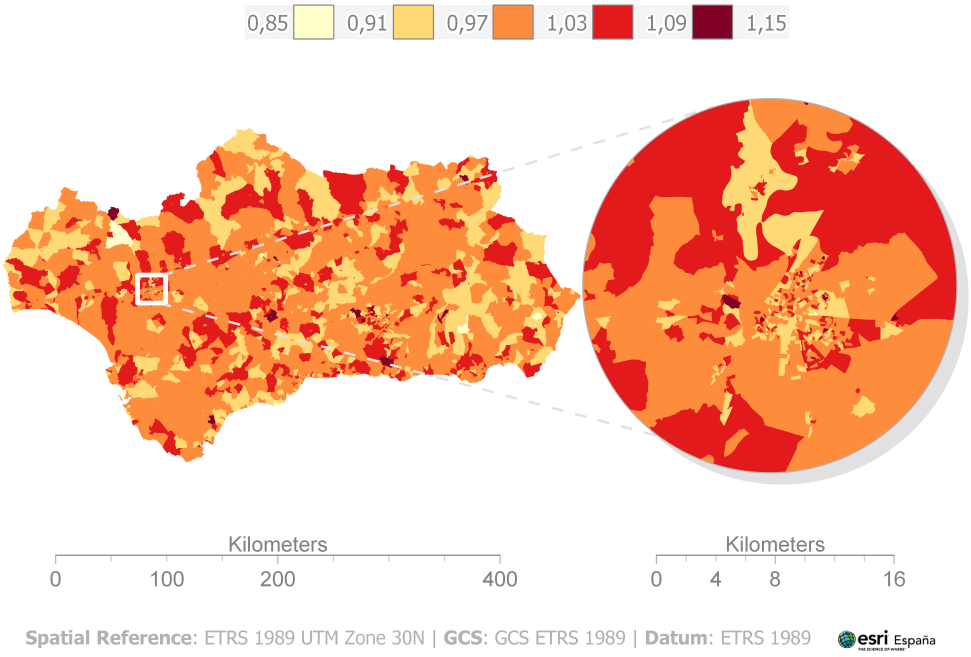


Fig. 4.5 Exponentiated random effects by census tract in Andalusia and Seville (zoom-in)

initial negative impact of "urbanicity" disappeared when incorporating other small area variables into the model, indicating that the physical urban concept may mask effects in other spheres. Although initial models appeared to show a small effect of the more precise physical measures on individual-level survival, the index does not explain small area differences in mortality in Andalusia after controlling for additional information on environmental and social measures. Because our results focus on a single region in Spain, these population and environment effects on survival should be examined in other contexts. Further analysis can highlight potential risk factors in different residential area types and their effect on growing inequalities in individual-level survival.

Some limitations and threats to validity exist due to data availability, unobserved mediators, and assumptions undertaken when conducting the analysis. Based on general life course trajectories in the context of Andalusia, we trust that individuals are unlikely to change their residence after age 35 [230]. However, while residency likely remains stable, other central explanatory variables such as our "urbanicity" indicator and some individual-level information have almost certainly changed over time. Since we cannot estimate the extent of these variations with the available single cross-sectional data point in 2001, we have to assume environmental and individual indicators like car ownership remain relatively constant over the time period of our study.

The most recent financial and debt crisis of 2008 hit Andalusia particularly hard and led to significant job loss, a continuously increasing at-risk-of-poverty rate, and a high number of evictions [90, 248]. Thus, the assumption of residential continuity may not hold among economically disadvantaged groups. Moreover, there is no information measuring the average exposure to the estimated environmental and social features of the residential area. The average amount of time someone spends in his or her residential area and is therefore exposed to its environment probably differs by age, employment status, and other unmeasured area features such as access to "third places" [213, 195].

In spite of these data constraints, to our knowledge, this analysis is the first study to combine detailed small area (census tract) information, individual-level variables, and survival follow up in Southern Europe. Our results differ from a previous analysis on a larger provincial scale for Spain [239]. This other study suggested a negative association between per-capita income and average survival times while we find that environmental features and, to a greater extent, population composition affect survival probability. Moreover, our exploratory survival analysis contributes to the debate on how individual health and socioeconomic differences relate to spatially correlated mortality differences. Such research can help to explain the mechanisms behind prevalent spatial-temporal inequalities such as those in Andalusia [207].

Although our results could not explain geographical differences based on a more detailed measure of urban space, we found that area conditions such as high levels of perceived pollution and a high percentage of unemployed co-residents increase individual relative mortality risks in the presence of other well-known protective individual characteristics. Future research must continue to explore and account for the role area heterogeneity plays in individual mortality.

Chapter 5

Conclusion

5.1 Contribution

Inequalities in health and survival appear to persists, regardless of collective actions and substantial investments in welfare and the goal to prevent such inequalities [cf. 64]. Measurable group differences in survival were found in all European countries independent from their welfare system or budget [180, 175]. There are countless pathways that might explain how socioeconomic and environmental features affect someone's health over the life course. In combination with the lack of universal measurement strategies for the assessment of overall health, such complexity requires creativity and caution when researching health inequalities. While creativity is necessary to break up complex relationships, it is important to be cautious because health care and social security systems are dependent on this kind of research. Imprecise analyses and projections about how such inequalities evolve in the future have the potential to cause substantial economic loss or, worse, the loss of lives or healthy life years.

In this thesis, I present research that quantifies existing inequalities in health and survival in and within Spain with the goal to contribute to a better understanding of future challenges regarding the potential effects of such inequalities. My co-authors and I shed a light on how different aspects of someone's social position and related external factors determine mortality at older ages. The three articles presented in this thesis touch on the potential effects of pension income, educational attainment, spatial inequalities, and disability pathways on old-age survival in Spain.

In the first article (chapter 2), the implications of retirement income inequality on survival are examined and related to the recent incorporation of average life expectancy in the

formula to determine the eligibility age for public pensions in Spain [cf. 255]. By comparing different income groups in Andalusia, the findings suggest striking survival disadvantage for the group with the lowest household income when compared with their wealthier counterparts. Although the effect is found for both men and women, it is substantially larger for male retirees.

The goal of the article presented in chapter 3 is to compare different pathways through disability and their effect on the survival in individuals aged 50 and older. By classifying disability trajectories, three general pathways were identified. Then survival difference by trajectory type and various socioeconomic factors are compared with their counterparts who do not suffer from any form of disability. While the results suggest that disability has an equalizing effect regarding the association of socioeconomic inequalities with mortality, the findings also indicate that the fatality of different trajectories strongly depend on the ability to become accustomed to a live with disability. Younger individuals are found to adapt better and survive longer even after the onset of severe disabilities.

The focus of chapter 4 is on how long-term exposure to environmental risks in rural and urban areas influence health and old-age survival. Although outshined by individual level characteristics the results showed significantly higher mortality risks for people who live in areas with greater levels of perceived pollution and unemployment. Furthermore, our results suggest that the population composition with regard to social position affects individual survival risks independent from their own social position. The degree of "urbanicity", how urban a neighborhood is, does not show any effects on survival when individual differences are controlled for.

Although the three articles may focus on three relatively different sets of impact factors and populations, the work is connected by an underlying interest. The shared goal of all three papers is to contribute to the better understanding of mechanisms that explain how socioeconomic position, environmental, and sociocultural factors across the life course affect disparities related to the onset of ill-health and survival. Branching out and touching on different impact factors allowed for capturing different dimensions of this big theoretical question and better contextualize these wide concepts such as health inequality. Furthermore, it opened different alleys for future research and my personal academic future.

The question of how socioeconomic position is related to later life health and survival was never more important than in times of unprecedented population aging. Large cohorts are

coming to ages where prevalence of chronic diseases and disability becomes more common. As the first baby boomers have reached retirement age, many Western countries expect an explosion of future health care costs [56, 143]. Even if health spans can be further extended, there will be an ever-larger population suffering from chronic conditions and non-communicable diseases such as dementia. At the same time a shrinking working age population will have to stem a growing demand of public pensions and care need. These developments make the understanding of how inequalities affect the onset of morbidity or average life lengths is indispensable. The quantification of risks and survival probabilities for different social groups can provide a foundation to plan ahead, assure a certain living standard, and improve social justice. Moreover, this research is important to better understand future public health needs, identify vulnerable groups, and prevent further social polarization.

It is essential to understand the mechanics behind health inequalities to assure that health services remain accessible for everyone. Recent reports from the USA, China, the UK, and other European countries indicate that already a large share of elderly in need of care is left unattended. This imposes large risks on this vulnerable population and their families. The word crisis is regularly used when the care situation is addressed publicly [cf. 71]. The privatization of the care sector is often proposed as solution for crammed nursing homes and the increasing burden for the public hospitals and doctors [194, 27]. While there is increasing evidence that privatization can complement public care and offer good quality care, proponents of this solution tend to ignore that structural inequalities will prevent especially the most vulnerable population from taking advantage of these private care options [132**?**].

The impending opening of the health gap also presents a challenge for law makers who will need to understand trends and risks to adapt social security policies or regulate the vastly growing care and pharma sector. The better we understand the risks and possible challenges, the earlier we can adapt and enact plans that will hopefully mitigate the effects of the looming health care crisis.

Better or new answers to the question how social position affects later-life health might also help us to better understand how the process that leads to an early onset of disease or an early death is related to different stages of the life course [290]. At an individual level exposure to structural disadvantages has been found to be associated to higher levels of stress, more relative exposure to judgement, and feelings of being left out. Such often hidden impact factors can be linked to a whole cascade of beliefs, behaviors, and habit loops that contributes to an early decline of an individuals' mental and physical health. The constant

constraints and feelings of inadequacy ultimately manifest itself in disease and disability [170, 28]. Although these relationships are difficult to analyze quantitatively, they can be accounted for as underlying mechanism and need to be better understood to make effective health policies.

In the Spanish context, it will be important to observe group differences in health and mortality in the future despite continued growth of life expectancy at birth and the relative equality in terms of healthy life spans [226]. Although the rapidly aging Spanish population was recently projected to become the forerunner in longevity by 2040, there appear to be lagged health effects that have not caught up yet to the rapid social, economic, and population development [98, 165]. For example, the latest financial and debt crisis hit Spain extraordinary hard. After the financial markets broke down, the artificially inflated construction sector collapsed, and many other sectors followed like dominoes. Unemployment rates skyrocketed and remained especially high for young people and those who perform forms of manual labor, who were left with the options to re-enter or stay in the education system or emigrate to larger cities or abroad [288, 131].

While those who did not emigrate were mostly able to fall back on family support systems, entry jobs were either hard to find or only short-term solutions [81, 84]. In turn an insider-outsider labor market emerged, and assets and wealth cumulated in the hands of an ever-smaller group. There were also those families and single parent households that were not able to support and protect their younger generation effectively. Many people, often members of the here identified vulnerable groups, lost their mortgage to the bank or had to file bankruptcy. Consequently, the at-risk poverty rates climbed to new highs [248]. Austerity measures in response to the crisis further let to cuts in health care and social expenditure funds [168]. Thus, since the last financial crisis many people, especially those on the lower rungs of the socioeconomic ladder, have faced problems to enter the relatively exclusive and only slowly recovering job market, they have been exposed to a crumbling health system, and many of them will not have the chance to save money during their presumably most productive years.

It is not imperative that the health gap will diverge in the future, but there is potential that the elderly of the future will be more unequal in terms of their socioeconomic backgrounds and consequentially their health. Although increasing economic inequalities have not yet manifested themselves in differences in physical diseases, increased prevalence of mental disorders and depression indicate a worrisome trend, as it is assumed that many of the

new cases are crisis and poverty risk related [110]. The assumptions that the relative health equality in Spain is likely to be jeopardized by the long-term effect of the latest crisis cannot be dismissed [see 165]. Therefore, there is the sustained need for comprehensive analyses and projections of inequalities in health and survival that focus on different aspects of social life.

5.2 Limitations

The applicability and meaningfulness of research output from life course studies and the analysis of health inequalities is dependent on multiple factors. It can be limited, for example, by the lack of longitudinal data, poor conceptualization, or the region-specific context. Among the limitations that have potentially affected the results and interpretations presented in this thesis are two reoccurring threats to validity. The first limitation, the repeated use of mortality as approximation for overall health, proposes a threat to the construct validity especially for the first and the third article [see 61].

While one could argue that at a population level deteriorating health is closely linked with temporal proximity to death, this relationship is strongly affected by shifts in the social structure and, presumably more important, medical progress [35]. If formerly deadly diseases can be diagnosed earlier, better treated, or even cured, there will be an effect on the health spans of people as for example discussed in chapter 3. Especially the progress regarding the detection and treatment of "silent diseases" such as cancer or diabetes type 2, led to the prolongation of the lives of more "unhealthy" individuals, who would have died without these new medical means [145]. These developments will mitigate the temporal correlation of diseases onset and death, and presumably limit the arguments for mortality as an indicator for a recent onset of ill-health. Although health span trends vary by region and by the definition of health itself, prevalence of chronic and age-related diseases is increasing in virtually all modern, aging societies [cf. 62, 56].

While these findings support the hypothesis that disease onset and death may become more detached temporarily, when using functionality measures and onset of care need, there tend to be a compression of the period between the onset of disease and death [224]. The use of only these two measurement strategies for assessing the same underlying construct can lead to very different interpretations for the same population. Thus, the use of such a wide concept comes with the advantage of relative flexibility but also the responsibility to guide the reader to understand what is not measured and which subpopulation are not included in the analysis. These conceptual problems also affect the measurement strategies. The onset ill-health is for example not universally experienced and has a variety of different implications for a person's life, depending on the severity and personal situation. Death, on the other hand, is an unambiguous state that everyone experiences. Therefore, it fulfills, among others, the underlying assumptions of time to event regression approaches like the Cox PH model that everyone is expected to experience the event of interest at some point in the future [9].

Furthermore, individual level health data is highly sensitive and for good reasons hard to access. With a few exceptions, like the population registers in the North European countries and the United Kingdom, I am not aware of many population-based, longitudinal databases on different aspects of health. In Spain, there are efforts to exploit regional register data bases to examine population health. The census based BDLPA data can potentially be linked to individual health data from Andalusia, which would allow for further explorations of health trends in South Europe. Unfortunately, these linkages are long-lasting operations with political implications and the collaboration in research groups requires numerous rounds of negotiations. At the time this thesis was written, the different parties have regrettably not reached an agreement yet.

The linked data, I had the opportunity to work with, was unique in a sense that these linkages between surveys or census data and a register-based follow-up have only recently been made available to social science research. Although they present numerous opportunities for answering new research questions in the realm of population health and mortality research, these data sources were not necessarily designed for longitudinal analyses. Some of the difficulties are touched upon in the articles presented in this thesis. To summarize, both (or technically all three) data sources are rich in individual level information and they allow to follow up survival over several years, 9 years in the case of the linkage of EDAD and up to 17 years for the BDLPA. However, both datasets are based on cross-sectional studies that were linked to longitudinal mortality information and the *Padrón* to account for changes in the exposure. Consequentially, individual level context information is time invariant. This may not be problematic when using covariates like highest educational attainment to explain mortality differences. Nonetheless, it can require one to make bold assumptions about other features as in chapter 4 where it is assumed that the residential area have remained the same in terms of composition and exposure to environmental threats.

Moreover, this kind of data requires you to use survival models that allow for left truncation. As explained above these survey populations are selected upon survival and the age structure may have been affected by unobserved period effects before the survey or census year. Since age is used as time dimension in our models, the populations of cross-sectional studies are comparable to a period life table population which resembles a hypothetical cohort.

5.3 Future Research

Population aging will challenge how we live in our modern societies. I have touched on some of these challenges throughout this thesis and sometimes the continuous warnings create the impression that a grim future is waiting for us. However, demographic aging is a success story and arguably one of the greatest achievements of humanity [243]. Medical and social progress within the last century has granted millions the opportunity to live up to old ages and remain a certain quality of life even after onset of disability.

To reliably predict future demographic scenarios and prepare for the challenges ahead, it will be essential to understand the interactions between social factors, environment, and the biological processes related to aging, onset of morbidity, and death. While expert knowledge will be necessary to understand the mechanics and physics of the aging process, the key to the sustainable generation of knowledge about these interactions will require to build more interdisciplinary research networks. The role of social scientists and demographers in this process is likely to change. They will have to adapt and expand their methodological toolbox. Personally, I see three promising research alleys for demographers and other social scientists who are interested in how health equality research is evolving.

First, there appears to be increased interest in sharing and analyzing administratively collected data. Information on disease and treatment are routinely collected by different private and public institutions, such as health insurance companies or hospitals. In some cases, these data points can be linked to other administratively collected pieces of information and allow for reconstructing individual life courses. Databases, like the BDPLA, are relatively cost-efficient complements to classic surveys and often capture information from large groups and sometimes full populations. Historical administrative data sources, as for example the Historical Sample of the Netherlands (HSN), further allow researcher to understand how past developments have affected present trends and investigate inter-generational transmission of social class. Moreover, we would be able to decompose to what extent health inequalities are produced by health selection and what part is attributed by the accumulation of disadvantages over the life course.

Second, recent advances in genetic epidemiology combined with new emerging computational tools for machine learning have the potential to change and improve the ways we investigate what is at the core of repeatedly observed effects of socio-cultural environment on later life health [229]. Some sociogenomic data sources, like the UK Biobank, include information on the genetic make-up or epigenetic markers related to the aging process and

the biology behind structural inequalities [cf. 23]. The knowledge about how such markers are transmitted and activated will complement classic nurture-based explanation of how socio-cultural factors determine health spans [cf. 162]. When such information is collected routinely, possibly complemented with health insurance data on treatments and medications, it will allow researchers to advance to new levels of understanding. We will be able to answer how the environment we live and grow up in influences biological responses of our body that then lead to measurable health outcomes.

When addressing the association between social conditions across the life course and later life health disparities, it will be necessary to better understand the role of mental health and how the emerging mental health crisis is related to demographic and societal shifts [161]. Presumably induced by the challenges of an ever-faster globalizing world, prevalence rates in mental disorders, such as burn-out and depression, are on the rise in most Western developed societies. Other than the majority of physical diseases, mental diseases do not only manifest themselves at the end of the life course but are common among adolescents and young adults, too [cf. 124, 153, 218]. Therefore, the economic impact due to hidden costs and loss of labor force may even be bigger than for many aging related diseases. In fact, mental diseases account for about a fifth of the overall burden of disease and are found to be precursors of various types of disability [159, 160].

Although mental disorders are hard to detect and can often not be diagnosed easily, a better understanding of how these diseases are associated with structural inequalities is essential to predict the future burden of disease and learn more about the connection of body and mind [7]. Apart from the general observations that individuals with lower social status are more likely to develop a mental disorder [166, 242], there is a need to further explore how demographic trends of family formation and migration to urban areas lead to changes in communities and environment that potentially enhance the exposure to known risk factors for mental diseases, such as stress and loneliness [cf. 80].

5.4 Policy recommendations

There are various potential lines of action that would assure a better detection of health and survival inequalities at population level. While life expectancy is well-understood, routinely calculated and, for example, used for the formula to determine the optimal pension age in Spain, policy makers are often less familiar with the interpretation of lifespan disparity measures, like the life table entropy, the concept life years lost due to death (e^{dagger}), or the standard deviation above the mode. Such measures may be more difficult to interpret but they could provide decision makers with equality trends that then help them to detect emerging problems at larger scale, such as the opioid crisis in the USA [293]. Moreover, they can be relatively easily derived from life tables that are already produced to obtain life expectancy.

High lifespan disparity is predominantly found among groups or within areas with low socioeconomic status. While low social status is associated with living in areas with high crime rates, high prevalence of drug abuse, and other factors correlated to high levels of young adult mortality, an increased uncertainty with regard to the timing of death reduces the chances for these people to break out of the vicious cycle. People are known to estimate their own life expectancy and chances to survive up to a certain age by observing the timing of deaths of their relatives, friends, and neighbors. Based on these estimation, they make decisions about investments, for example, in education, which then affects their later life affluence and ultimately their health [259]. To break through this cycle and make savings and investments to old age more worthwhile, a general, measurable goal could be to reduce overall and with-in group disparities in average lifespans [296].

First groups with high lifespan and healthy lifespan disparities need to be identified. The second step is then to analyze their specific context to find ways to break out of the downward spiral. It will be necessary to determine when behaviors and actions that lead to poor health and early deaths are generally adopted and how predispositions are transmitted through the genes of previous generations. It will be important to disentangle what part of observed inequalities exist because of selection processes based inherited or early obtained bad health and what part is accumulated over the life course. Moreover, it will be necessary to identify what environmental risks need to be reduced to assure a certain quality of living.

To conclude, inequalities in health and survival are at the core of every other form of inequality due to their functioning in selection processes or as reason for failed investments. Their reduction is a social justice and human rights issue that must have highest priority on all political agendas. While the urgency of an opening health and life expectancy gap ask for

immediate action, the complexity of the relationship between our environment, our social position, and health requires extreme caution when drawing inferences from research. It will be essential to continue efforts to build interdisciplinary research groups and combine expert knowledge to comprehend how the social and genetic predisposition are transformed over the life course and determine how we age.

References

[1] 8/2002, L. (2002). Ley 8/2002, de 17 de diciembre, del plan estadístico de Andalucía 2003-2006, y de modificación de la ley 4/1989, de 12 de diciembre, de estadística de la comunidad autónoma de Andalucía.

[2] Abbott, A. and Tsay, A. (2000). Sequence analysis and optimal matching methods in sociology: Review and prospect. *Sociological methods & research*, 29(1):3–33.

[3] Abellán García, A., Pérez Díaz, J., Esparza Catalán, C., Castejón Villarejo, P., and Ahmed Mohamed, K. (2012). *Discapacidad y dependencia en Andalucía. Con base en la Encuesta sobre Discapacidad, Autonomìa personal y situaciones de Dependencia (EDAD) 2008*. Andalucía. Junta, Seville.

[4] Abellán, A., Esparza, C., Castejón, P., and Pérez, J. (2011). Epidemiología de la discapacidad y la dependencia de la vejez en españa. *Gaceta Sanitaria*, 25:5–11.

[5] Abubakar, I., Tillmann, T., and Banerjee, A. (2015). Global, regional, and national age-sex specific all-cause and cause-specific mortality for 240 causes of death, 1990-2013: a systematic analysis for the global burden of disease study 2013. *Lancet*, 385(9963):117–171.

[6] Adams, P., Hurd, M. D., McFadden, D., Merrill, A., and Ribeiro, T. (2003). Healthy, wealthy, and wise? tests for direct causal paths between health and socioeconomic status. *Journal of econometrics*, 112(1):3–56.

[7] Allen, J., Balfour, R., Bell, R., and Marmot, M. (2014). Social determinants of mental health. *International review of psychiatry*, 26(4):392–407.

[8] Allender, S., Lacey, B., Webster, P., Rayner, M., Deepa, M., Scarborough, P., Arambepola, C., Datta, M., and Mohan, V. (2010). Level of urbanization and noncommunicable disease risk factors in Tamil Nadu, India. *Bulletin of the World Health Organization*, 88(4):297–304.

[9] Allison, P. D. (2010). *Survival analysis using SAS: a practical guide*. Sas Institute.

[10] Allison, P. D. (2014). *Event history and survival analysis: Regression for longitudinal event data*, volume 46. SAGE publications.

[11] Altindag, D., Cannonier, C., and Mocan, N. (2011). The impact of education on health knowledge. *Economics of Education Review*, 30(5):792–812.

[12] Anderson, P. and del Arco Blanco, M. Á. (2014). *Mass Killings and Violence in Spain, 1936-1952: Grappling with the past*. Routledge.

[13] Anselin, L. (2013). *Spatial econometrics: methods and models*, volume 4. Springer Science & Business Media.

[14] Anselin, L., Syabri, I., and Smirnov, O. (2002). Visualizing multivariate spatial correlation with dynamically linked windows. *Urbana*, 51:61801.

[15] Asián-Chaves, R. (2000). ¿ terciarización de la economía andaluza? la estructura productiva andaluza y los servicios en la globalización. *Revista de estudios regionales, 58, 79-111*.

[16] Austin, P. C. (2017). A tutorial on multilevel survival analysis: methods, models and applications. *International Statistical Review*, 85(2):185–203.

[17] Balboa-Castillo, T., Guallar-Castillón, P., León-Muñoz, L. M., Graciani, A., López-García, E., and Rodríguez-Artalejo, F. (2011). Physical activity and mortality related to obesity and functional status in older adults in spain. *American journal of preventive medicine*, 40(1):39–46.

[18] Baltes, M. M. (1996). *The many faces of dependency in old age*. Cambridge University Press.

[19] Bauer, U. E., Briss, P. A., Goodman, R. A., and Bowman, B. A. (2014). Prevention of chronic disease in the 21st century: elimination of the leading preventable causes of premature death and disability in the USA. *The Lancet*, 384(9937):45 – 52.

[20] Bauman, A., Merom, D., Bull, F. C., Buchner, D. M., and Fiatarone Singh, M. A. (2016). Updating the evidence for physical activity: summative reviews of the epidemiological evidence, prevalence, and interventions to promote "active aging". *The gerontologist*, 56(Suppl_2):S268–S280.

[21] Beehr, T. A. and Bowling, N. A. (2013). Variations on a retirement theme: Conceptual and operational definitions of retirement. *The Oxford Handbook of Retirement*.

[22] Belloni, M., Alessie, R., Kalwij, A., and Marinacci, C. (2013). Lifetime income and old age mortality risk in Italy over two decades. *Demographic Research*, 29:1261–1298.

[23] Belsky, D. W., Domingue, B. W., Wedow, R., Arseneault, L., Boardman, J. D., Caspi, A., Conley, D., Fletcher, J. M., Freese, J., Herd, P., et al. (2018). Genetic analysis of social-class mobility in five longitudinal studies. *Proceedings of the National Academy of Sciences*, 115(31):E7275–E7284.

[24] Beltrán-Sánchez, H., Razak, F., and Subramanian, S. (2014). Going beyond the disability-based morbidity definition in the compression of morbidity framework. *Global health action*, 7(1):24766.

[25] Bengtson, V. L., Elder, G. H., and Putney, N. M. (2005). *The Lifecourse Perspective on Ageing: Linked Lives, Timing, and History*, pages 493–501. Cambridge Handbooks in Psychology. Cambridge University Press, Cambridge.

[26] Åberg, A. C., Sidenvall, B., Hepworth, M., O'Reilly, K., and Lithell, H. (2005). On loss of activity and independence, adaptation improves life satisfaction in old age – a qualitative study of patients' perceptions. *Quality of Life Research*, 14(4):1111–1125.

[27] Bergman, M. A., Johansson, P., Lundberg, S., and Spagnolo, G. (2016). Privatization and quality: Evidence from elderly care in sweden. *Journal of Health Economics*, 49:109–119.

[28] Berkman, L. F., Kawachi, I., and Glymour, M. M. (2014). *Social epidemiology*. Oxford University Press.

[29] Bermudez, A. and Brey, E. (2017). *Is Spain Becoming a Country of Emigration Again? Data Evidence and Public Responses*, pages 83–98. Springer.

[30] Bideau, A. and Brignoli, H. P. (1997). *Infant and child mortality in the past*. Oxford University Press.

[31] Billari, F. C. and Piccarreta, R. (2005). Analyzing demographic life courses through sequence analysis. *Mathematical population studies*, 12(2):81–106.

[32] Bivand, R. S., Pebesma, E., and Gómez-Rubio, V. (2013). *Applied Spatial Data Analysis with R*, volume 10. Springer Science & Business Media.

[33] Bloemen, H., Hochguertel, S., and Zweerink, J. (2017). The causal effect of retirement on mortality: Evidence from targeted incentives to retire early. *Health Economics*.

[34] Boldrin, M. and Jiménez-Martín, S. (2007). *Evaluating Spanish pension expenditure under alternative reform scenarios*, pages 351–412. University of Chicago Press.

[35] Bowling, A. (2004). *Measuring health*. McGraw-Hill Education (UK).

[36] Braveman, P. (2014). What are health disparities and health equity? we need to be clear. *Public health reports (Washington, D.C. : 1974)*, 129 Suppl 2(Suppl 2):5–8.

[37] Braveman, P. and Barclay, C. (2009). Health disparities beginning in childhood: a life-course perspective. *Pediatrics*, 124(Supplement 3):S163–S175.

[38] Braveman, P., Egerter, S., and Williams, D. R. (2011). The social determinants of health: coming of age. *Annual review of public health*, 32:381–398.

[39] Breen, R. and Breen Jr, R. (2004). *Social mobility in Europe*. Oxford University Press.

[40] Brown, D. C., Hayward, M. D., Montez, J. K., Hummer, R. A., Chiu, C.-T., and Hidajat, M. M. (2012). The significance of education for mortality compression in the united states. *Demography*, 49(3):819–840.

[41] Brown, T. A. (2014). *Confirmatory factor analysis for applied research*. Guilford Publications.

[42] Brzinsky-Fay, C. and Kohler, U. (2010). New developments in sequence analysis.

[43] Buchman, A. S., Wilson, R. S., Yu, L., James, B. D., Boyle, P. A., and Bennett, D. A. (2014). Total daily activity declines more rapidly with increasing age in older adults. *Archives of gerontology and geriatrics*, 58(1):74–79.

[44] Cain, K. C., Harlow, S. D., Little, R. J., Nan, B., Yosef, M., Taffe, J. R., and Elliott, M. R. (2011). Bias due to left truncation and left censoring in longitudinal studies of developmental and disease processes. *American journal of epidemiology*, 173(9):1078–1084.

[45] Calero, D. and Navarro, E. (2011). Differences in cognitive performance, level of dependency and quality of life (qol), related to age and cognitive status in a sample of spanish old adults under and over 80 years of age. *Archives of Gerontology and Geriatrics*, 53(3):292–297.

[46] Camarda, C. G. (2012). Mortalitysmooth: An r package for smoothing poisson counts with p-splines. *Journal of Statistical Software*, 50(1):1–24.

[47] Cambois, E., Blachier, A., and Robine, J.-M. (2013). Aging and health in France: an unexpected expansion of disability in mid-adulthood over recent years. *European Journal of Public Health*, 23(4):575–581.

[48] Canchola, A. J., Stewart, S. L., Bernstein, L., West, D. W., Ross, R. K., Deapen, D., Pinder, R., Reynolds, P., Wright, W., and Anton-Culver, H. (2003a). Cox regression using different time-scales. *Western Users of SAS Software. San Francisco, California*.

[49] Canchola, A. J., Stewart, S. L., Bernstein, L., West, D. W., Ross, R. K., Deapen, D., Pinder, R., Reynolds, P., Wright, W., and Anton-Culver, H. (2003b). Cox regression using different time-scales. *Western Users of SAS Software. San Francisco, California*.

[50] Canudas-Romo, V. (2008). The modal age at death and the shifting mortality hypothesis. *Demographic Research*, 19:1179–1204.

[51] Casanova, J., Maestre, F. E., Mir, C., and Gómez, F. M. (2002). *Morir, matar, sobrevivir: La violencia en la dictadura de Franco*. Crítica.

[52] Caselli, G. (2010). Increasing longevity and decreasing gender mortality differentials: new perspectives from a study on Italian cohorts. *Women*, 65(e65M):e65M.

[53] Caspi, A., Harrington, H., Moffitt, T. E., Milne, B. J., and Poulton, R. (2006). Socially isolated children 20 years later: risk of cardiovascular disease. *Archives of pediatrics & adolescent medicine*, 160(8):805–811.

[54] Cheng, G. and Zervopoulos, P. D. (2014). Estimating the technical efficiency of health care systems: A cross-country comparison using the directional distance function. *European Journal of Operational Research*, 238(3):899–910.

[55] Chetty, R., Stepner, M., Abraham, S., Lin, S., Scuderi, B., Turner, N., Bergeron, A., and Cutler, D. (2016). The association between income and life expectancy in the united states, 2001-2014. *Jama*, 315(16):1750–1766.

[56] Christensen, K., Doblhammer, G., Rau, R., and Vaupel, J. W. (2009). Ageing populations: the challenges ahead. *The lancet*, 374(9696):1196–1208.

[57] Cockerham, W. C. (2005). Health lifestyle theory and the convergence of agency and structure. *Journal of health and social behavior*, 46(1):51–67.

[58] Cohen, J. E., Bohk-Ewald, C., and Rau, R. (2018). Gompertz, makeham, and siler models explain taylor's law in human mortality data. *Demographic Research*, 38:773–841.

[59] Colchero, F., Rau, R., Jones, O. R., Barthold, J. A., Conde, D. A., Lenart, A., Nemeth, L., Scheuerlein, A., Schoeley, J., Torres, C., Zarulli, V., Altmann, J., Brockman, D. K., Bronikowski, A. M., Fedigan, L. M., Pusey, A. E., Stoinski, T. S., Strier, K. B., Baudisch, A., Alberts, S. C., and Vaupel, J. W. (2016). The emergence of longevous populations. *Proceedings of the National Academy of Sciences*, 113(48):E7681–E7690.

[60] Colvin, V. L. (2003). The potential environmental impact of engineered nanomaterials. *Nature biotechnology*, 21(10):1166.

[61] Creswell, J. W. and Creswell, J. D. (2017). *Research design: Qualitative, quantitative, and mixed methods approaches*. Sage publications.

[62] Crimmins, E. M., Alley, D., Reynolds, S. L., Johnston, M., Karlamangla, A., and Seeman, T. (2005). Changes in biological markers of health: Older americans in the 1990s. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, 60(11):1409–1413.

[63] Crimmins, E. M., Zhang, Y., and Saito, Y. (2016). Trends over 4 decades in disability-free life expectancy in the United States. *American journal of public health*, 106(7):1287–1293.

[64] Crombie, I., Irwine, L., Elliot, L., and Wallace, H. (2005). Closing the health inequalities gap: An international perspective.

[65] CSDH (2008). Closing the gap in a generation: health equity through action on the social determinants of health. Report, World Health Organization. Final Report of the Commission on Social Determinants of Health.

[66] Cutler, D. M. and Lleras-Muney, A. (2010). Understanding differences in health behaviors by education. *Journal of health economics*, 29(1):1–28.

[67] Cylus, J. and Pearson, M. (2016). Cross-national efficiency comparisons of health systems, subsectors and disease areas. *Health System Efficiency*, page 139.

[68] Cyril, S., Oldroyd, J. C., and Renzaho, A. (2013). Urbanisation, urbanicity, and health: a systematic review of the reliability and validity of urbanicity scales. *BMC Public Health*, 13(1):513.

[69] Dahl, E. and van der Wel, K. A. (2013). Educational inequalities in health in european welfare states: a social expenditure approach. *Social Science & Medicine*, 81:60–69.

[70] Dahly, D. L. and Adair, L. S. (2007). Quantifying the urban environment: A scale measure of urbanicity outperforms the urban–rural dichotomy. *Social Science & Medicine*, 64(7):1407–1419.

[71] Dann, T. (2014). Global elderly care in crisis. *The Lancet*, 383(9921):927.

[72] Dave, D., Rashad, I., and Spasojevic, J. (2006). The effects of retirement on physical and mental health outcomes. Report, National Bureau of Economic Research.

[73] Díaz-Giménez, J. and Díaz-Saavedra, J. (2009). Delaying retirement in Spain. *Review of Economic dynamics*, 12(1):147–167.

[74] De la Porte, C. and Natali, D. (2014). Altered europeanisation of pension reform in the context of the great recession: Denmark and italy compared. *West European Politics*, 37(4):732–749.

[75] Deindl, C., Brandt, M., and Hank, K. (2016). Social networks, social cohesion, and later-life health. *Social Indicators Research*, 126(3):1175–1187.

[76] Del Boca, D. and Locatelli, M. (2003). Fertility and women's work status. *The Rationale of Motherhood Choices: influences of employment conditions and of public policies, European Commission, Belgium*, pages 149–169.

[77] del Campo Urbano, S. and López, M. N. (1987). *Nuevo análisis de la población española*. Ariel.

[78] Demakakos, P., Biddulph, J. P., Bobak, M., and Marmot, M. G. (2015). Wealth and mortality at older ages: a prospective cohort study. *Journal of Epidemiology and Community Health*, pages jech–2015–206173.

[79] Díaz-Salazar, G. L. (1999). La dinámica de la población en andalucía: transición y cambios en el siglo xx. *Boletín económico de Andalucía*, (25):129–152.

[80] Dickens, A. P., Richards, S. H., Greaves, C. J., and Campbell, J. L. (2011). Interventions targeting social isolation in older people: a systematic review. *BMC public health*, 11(1):647.

[81] Dolado, J. J., Jansen, M., Felgueroso, F., Fuentes, A., and Wölfl, A. (2013). Youth labour market performance in spain and its determinants. Technical Report 1039.

[82] Domínguez-Mujica, J., Díaz-Hernández, R., and PArrEño-cAstELLAno, J. (2016). Migrating abroad to get ahead: the emigration of young spanish adults during the financial crisis (2008–2013). In *Global change and human mobility*, pages 203–223. Springer.

[83] Dong, X., Milholland, B., and Vijg, J. (2016). Evidence for a limit to human lifespan. *Nature*, 538:257.

[84] Dudel, C., Gómez, M. A. L., Benavides, F. G., Myrskylä, M., et al. (2016). *A Lost Generation?: The Financial Crisis and the Length of Working Life in Spain*. Max-Planck-Institut für demografische Forschung.

[85] Eckert, S. and Kohler, S. (2014). Urbanization and health in developing countries: a systematic review. *World Health Popul*, 15(1):7–20.

[86] Eibich, P. (2015). Understanding the effect of retirement on health: mechanisms and heterogeneity. *Journal of health economics*, 43:1–12.

[87] Elder, G. H. (1998). The life course as developmental theory. *Child Development*, 69(1):1–12.

[88] Escolar-Pujolar, A., Bacigalupe, A., and San Sebastian, M. (2014). European economic crisis and health inequities: research challenges in an uncertain scenario. *International Journal for Equity in Health*, 13(1):59.

[89] Espenshade, T. J. and Tannen, J. B. (2015). Population dynamics: Momentum of population growth. In Wright, J. D., editor, *International Encyclopedia of the Social & Behavioral Sciences*, volume 6. Elsevier, 2 edition.

[90] Eurostat (2017a). European Union statistics on income and living conditions. Technical report, European Union.

[91] Eurostat (2017b). Keyfigures on Europe - 2016 edition. Report, European Union.

[92] Fergusson, J. E. (1990). *Heavy elements: chemistry, environmental impact and health effects*. Pergamon.

[93] Fernández-Tabales, A. and Cruz, E. (2013). Análisis territorial del crecimiento y la crisis del sector de la construcción en españa y la comunidad autónoma de andalucía. *EURE (Santiago)*, 39(116):5–37.

[94] Ferrucci, L., WINDHAM, B. G., and FRIED, L. P. (2005). Frailty in older persons. *Genus*, pages 39–53.

[95] Fiorillo, D. and Sabatini, F. (2011). Quality and quantity: the role of social interactions in self-reported individual health. *Social Science & Medicine*, 73(11):1644–1652.

[96] Fisher, G. G., Chaffee, D. S., and Sonnega, A. (2016). Retirement timing: A review and recommendations for future research. *Work, Aging and Retirement*, 2(2):230–261.

[97] for Economic Co-operation, O. and Development (2013). *OECD framework for statistics on the distribution of household income, consumption and wealth*. OECD Publishing.

[98] Foreman, K. J., Marquez, N., Dolgert, A., Fukutaki, K., Fullman, N., McGaughey, M., Pletcher, M. A., Smith, A. E., Tang, K., Yuan, C.-W., Brown, J. C., Friedman, J., He, J., Heuton, K. R., Holmberg, M., Patel, D. J., Reidy, P., Carter, A., Cercy, K., Chapin, A., Douwes-Schultz, D., Frank, T., Goettsch, F., Liu, P. Y., Nandakumar, V., Reitsma, M. B., Reuter, V., Sadat, N., Sorensen, R. J. D., Srinivasan, V., Updike, R. L., York, H., Lopez, A. D., Lozano, R., Lim, S. S., Mokdad, A. H., Vollset, S. E., and Murray, C. J. L. (2018). Forecasting life expectancy, years of life lost, and all-cause and cause-specific mortality for 250 causes of death: reference and alternative scenarios for 2016–40 for 195 countries and territories. *The Lancet*, 392(10159):2052–2090.

[99] Freedman, V. A., Martin, L. G., and Schoeni, R. F. (2002). Recent trends in disability and functioning among older adults in the united states: a systematic review. *Jama*, 288(24):3137–3146.

[100] Freudenberg, N., Galea, S., and Vlahov, D. (2005). Beyond urban penalty and urban sprawl: back to living conditions as the focus of urban health. *Journal of community health*, 30(1):1–11.

[101] Fried, L. P., Ferrucci, L., Darer, J., Williamson, J. D., and Anderson, G. (2004). Untangling the Concepts of Disability, Frailty, and Comorbidity: Implications for Improved Targeting and Care. *The Journals of Gerontology: Series A*, 59(3):M255–M263.

[102] Fries, J. F. (1980). Aging, natural death, and the compression of morbidity. *New England Journal of Medicine*, 303(3):130–135.

[103] Fries, J. F., Bruce, B., and Chakravarty, E. (2011). Compression of morbidity 1980–2011: a focused review of paradigms and progress. *Journal of aging research*, 2011.

[104] Gabadinho, A., Ritschard, G., Mueller, N. S., and Studer, M. (2011). Analyzing and visualizing state sequences in r with traminer. *Journal of Statistical Software*, 40(4):1–37.

[105] Galenkamp, H., Braam, A. W., Huisman, M., and Deeg, D. J. (2012). Seventeen-year time trend in poor self-rated health in older adults: changing contributions of chronic diseases and disability. *The European Journal of Public Health*, 23(3):511–517.

[106] Gallo, P. and Gené-Badia, J. (2013). Cuts drive health system reforms in spain. *Health Policy*, 113(1-2):1–7.

[107] García-Gómez, P., Jiménez-Martín, S., and Castelló, J. V. (2012). *Health, Disability, and Pathways into Retirement in Spain*, pages 127–174. University of Chicago Press.

[108] Gardner, J. W. and Sanborn, J. S. (1990). Years of potential life lost (ypll)—what does it measure? *Epidemiology*, pages 322–329.

[109] Gènova-Maleras, R., Catalá-López, F., de Larrea-Baz, N. F., Álvarez Martín, E., and Morant-Ginestar, C. (2011). The burden of premature mortality in Spain using standard expected years of life lost: a population-based study. *BMC public health*, 11(1):787.

[110] Gili, M., Roca, M., Basu, S., McKee, M., and Stuckler, D. (2012). The mental health risks of economic crisis in Spain: evidence from primary care centres, 2006 and 2010. *European Journal of Public Health*, 23(1):103–108.

[111] Gill, T. M., Allore, H., Hardy, S. E., Holford, T. R., and Han, L. (2005). Estimates of active and disabled life expectancy based on different assessment intervals. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, 60(8):1013–1016.

[112] Gill, T. M., Gahbauer, E. A., Han, L., and Allore, H. G. (2010). Trajectories of disability in the last year of life. *New England Journal of Medicine*, 362(13):1173–1180.

[113] Gispert, R., de Arán Barés, M., and Puigdefàbregas, A. (2006). La mortalidad evitable: lista de consenso para la actualización del indicador en españa. *Gaceta Sanitaria*, 20(3):184–193.

[114] Gjonça, A., Tomassini, C., Toson, B., and Smallwood, S. (2005). Sex differences in mortality, a comparison of the United Kingdom and other developed countries. *Health Statistics Quarterly*, 26(2):6–16.

[115] Glei, D., Gómez Redondo, R., Argüeso, A., and Canudas-Romo, V. (2006). About mortality data for spain. *Human Mortality Database*.

[116] Gong, P., Liang, S., Carlton, E. J., Jiang, Q., Wu, J., Wang, L., and Remais, J. V. (2012). Urbanisation and health in China. *The Lancet*, 379(9818):843–852.

[117] González, M. J., Jurado, T., and Naldini, M. (2014). *Gender inequalities in southern Europe: woman, work and welfare in the 1990s*. Routledge.

[118] Graham, H. (2009). *Understanding health inequalities*. McGraw-Hill Education (UK).

[119] Gravelle, H. and Sutton, M. (2009). Income, relative income, and self-reported health in Britain 1979 - 2000. *Health Economics*, 18(2):125–145.

[120] Green, M. A., Subramanian, S., Vickers, D., and Dorling, D. (2015). Internal migration, area effects and health: Does where you move to impact upon your health? *Social Science & Medicine*, 136:27–34.

[121] Gu, D., Yang, F., and Sautter, J. (2016). Socioeconomic status as a moderator between frailty and mortality at old ages. *BMC geriatrics*, 16(1):151.

[122] Guan, P., Zheng, X., Chung, K., and Zhong, N. (2016). Impact of air pollution on the burden of chronic respiratory diseases in China: Time for urgent action. *The Lancet*, 388(10054):1939–1951.

[123] Guner, N., Kaya, E., and Sánchez-Marcos, V. (2014). Gender gaps in Spain: policies and outcomes over the last three decades. *SERIEs - Journal of the Spanish Economic Association*, 5(1):61–103.

[124] Gunnell, D., Kidger, J., and Elvidge, H. (2018). Adolescent mental health in crisis.

[125] Halpin, B. (2010). Optimal matching analysis and life-course data: The importance of duration. *Sociological Methods & Research*, 38(3):365–388.

[126] Harper, S. (2014). Economic and social implications of aging societies. *Science*, 346(6209):587–591.

[127] Hayward, M. D., Hummer, R. A., and Sasson, I. (2015). Trends and group differences in the association between educational attainment and u.s. adult mortality: Implications for understanding education's causal influence. *Social Science & Medicine*, 127:8–18.

[128] Herce, J., Arévalo, P., Alonso, P., Bravo, L., Cabrerizo, C., and Rodrígez, I. (2007). La economía de andalucía:diagnóstico estratégico.

[129] Hernaes, E., Markussen, S., Piggott, J., and Vestad, O. L. (2013). Does retirement age impact mortality? *Journal of health economics*, 32(3):586–598.

[130] Hernández de Cos, P., Jimeno, J. F., and Ramos, R. (2017). The spanish public pension system: current situation, challenges and reform alternatives. Occasional Papers 1701, Banco de España.

[131] Hernández Peinado, M., Montero González, B., et al. (2013). Un enfoque regional de la relación de los jóvenes con el mercado de trabajo.

[132] Horton, R. and Clark, S. (2016). The perils and possibilities of the private health sector. *Lancet (London, England)*, 388(10044):540.

[133] Hosmer, D., Lemeshow, S., and May, S. (2011). *Applied Survival Analysis: Regression Modeling of Time to Event Data: Second Edition*. Wiley Blackwell.

[134] Human Mortality Database, University of California, Berkeley (USA), and Max Planck Institute for Demographic Research (Germany) (2016). Available at www.mortality.org or www.humanmortality.de (data downloaded on 26.01.2017).

[135] Hurd, M. D. and McGarry, K. (2002). The predictive validity of subjective probabilities of survival. *The Economic Journal*, 112(482):966–985.

[136] Hurd, M. D., Smith, J. P., and Zissimopoulos, J. M. (2004). The effects of subjective survival on retirement and social security claiming. *Journal of Applied Econometrics*, 19(6):761–775.

[137] IECA (2016). Estadísticas Longitudinales de Supervivencia y Longevidad en Andalucía 2002-2013. Accessed: 09/03/2017.

[138] Imai, K. and Soneji, S. (2007). On the estimation of disability-free life expectancy: Sullivan's method and its extension. *Journal of the American Statistical Association*, 102(480):1199–1211.

[139] INE, I. N. d. E. (2010). Encuesta sobre Discapacidad, Autonomía personal y Situaciones de Dependencia (EDAD) - Metodología. Report, Instituto Nacional de Estadística.

[140] INE, I. N. d. E. (2016). Cifras de población a 1 de julio de 2016. Report, Instituto Nacional de Estadística.

[141] INE, I. N. d. E. (2017). Series detalladas desde 2002 -resultados por comunidades autonomas.

[142] Jackson, G. (2010). *La República española y la guerra civil, 1931-1939*. Barcelona: Crítica, 1990.

[143] Jagger, C., Weston, C., Cambois, E., Van Oyen, H., Nusselder, W., Doblhammer, G., Rychtarikova, J., Robine, J.-M., et al. (2011). Inequalities in health expectancies at older ages in the european union: findings from the survey of health and retirement in europe (share). *J Epidemiol Community Health*, 65(11):1030–1035.

[144] Jargowsky, P. A. (1996). Take the money and run: Economic segregation in us metropolitan areas. *American sociological review*, pages 984–998.

[145] Jeune, B. and Brønnum-Hansen, H. (2008). Trends in health expectancy at age 65 for various health indicators, 1987–2005, denmark. *European Journal of Ageing*, 5(4):279.

[146] Johnston, D. W. and Lee, W.-S. (2009). Retiring to the good life? the short-term effects of retirement on health. *Economics Letters*, 103(1):8–11.

[147] Jones-Smith, J. and Popkin, B. M. (2010). Understanding community context and adult health changes in China: development of an urbanicity scale. *Social science & medicine*, 71(8):1436–1446.

[148] Jöreskog, K. G. (1967). A general approach to confirmatory maximum likelihood factor analysis. *ETS Research Report Series*, 1967(2):183–202.

[149] Jürges, H., Reinhold, S., and Salm, M. (2011). Does schooling affect health behavior? evidence from the educational expansion in western germany. *Economics of Education Review*, 30(5):862–872.

[150] Kalugina, E., Radtchenko, N., and Sofer, C. (2009). How do spouses share their full income? identification of the sharing rule using self-reported income. *Review of Income and Wealth*, 55(2):360–391.

[151] Kannisto, V. (2001). Mode and dispersion of the length of life. *Population: An English Selection*, pages 159–171.

[152] Karanikolos, M., Mladovsky, P., Cylus, J., Thomson, S., Basu, S., Stuckler, D., Mackenbach, J. P., and McKee, M. (2013). Financial crisis, austerity, and health in europe. *The Lancet*, 381(9874):1323–1331.

[153] Karel, M. J., Gatz, M., and Smyer, M. A. (2012). Aging and mental health in the decade ahead: What psychologists need to know. *American Psychologist*, 67(3):184.

[154] Karlsdotter, K., Martín Martín, J. J., and López del Amo González, M. P. (2012). Multilevel analysis of income, income inequalities and health in Spain. *Social Science & Medicine*, 74(7):1099–1106.

[155] Katz, S. (1983). Assessing self-maintenance: activities of daily living, mobility, and instrumental activities of daily living. *Journal of the American Geriatrics Society*, 31(12):721–727.

[156] Kleinbaum, D. G. and Klein, M. (2010). *Survival analysis*, volume 3. Springer.

[157] Klijs, B., Mackenbach, J. P., and Kunst, A. E. (2010). Disability occurrence and proximity to death. *Disability and rehabilitation*, 32(21):1733–1741.

[158] Kluge, F., Zagheni, E., Loichinger, E., and Vogt, T. (2014). The advantages of demographic change after the wave: Fewer and older, but healthier, greener, and more productive? *PLOS ONE*, 9(9):e108501.

[159] Knapp, M. (2003). Hidden costs of mental illness. *The British Journal of Psychiatry*, 183(6):477–478.

[160] Knapp, M., McDaid, D., and Parsonage, M. (2011). Mental health promotion and mental illness prevention: The economic case. Technical report.

[161] Koenig, H. G., George, L. K., and Schneider, R. (1994). Mental health care for older adults in the year 2020: A dangerous and avoided topic. *The Gerontologist*, 34(5):674–679.

[162] Kong, A., Thorleifsson, G., Frigge, M. L., Vilhjalmsson, B. J., Young, A. I., Thorgeirsson, T. E., Benonisdottir, S., Oddsson, A., Halldorsson, B. V., Masson, G., et al. (2018). The nature of nurture: Effects of parental genotypes. *Science*, 359(6374):424–428.

[163] Kordel, S. (2016). The production of spaces of the 'good life' – the case of lifestyle migrants in Spain. *Leisure Studies*, 35(2):129–140.

[164] Krumpal, I. (2013). Determinants of social desirability bias in sensitive surveys: a literature review. *Quality & Quantity*, 47(4):2025–2047.

[165] Kulhánová, I., Bacigalupe, A., Eikemo, T. A., Borrell, C., Regidor, E., Esnaola, S., and Mackenbach, J. P. (2014). Why does Spain have smaller inequalities in mortality? an exploration of potential explanations. *European Journal of Public Health*, 24(3):370–377.

[166] Lai, E. T., Wickham, S., Law, C., Whitehead, M., Barr, B., and Taylor-Robinson, D. (2019). Poverty dynamics and health in late childhood in the uk: evidence from the millennium cohort study. *Archives of disease in childhood*, pages archdischild–2018.

[167] Lawton, M. P. and Brody, E. M. (1969). Assessment of older people: self-maintaining and instrumental activities of daily living. *The gerontologist*, 9(3):179–186.

[168] Legido-Quigley, H., Otero, L., la Parra, D., Alvarez-Dardet, C., Martin-Moreno, J. M., and McKee, M. (2013). Will austerity cuts dismantle the Spanish healthcare system? *BMJ*, 346.

[169] Leon, D. A. (2008). Cities, urbanization and health. *International Journal of Epidemiology*, 37(1):4–8.

[170] Link, B. G. and Phelan, J. (1995). Social conditions as fundamental causes of disease. *Journal of health and social behavior*, pages 80–94.

[171] Lipsitz, L. A. (2002). Dynamics of stability: the physiologic basis of functional health and frailty. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, 57(3):B115–B125.

[172] Lochner, L. (2011). Non-production benefits of education: Crime, health, and good citizenship. Report, National Bureau of Economic Research.

[173] Lutz, W., Sanderson, W., and Scherbov, S. (2008). The coming acceleration of global population ageing. *Nature*, 451(7179):716.

[174] Lynn, R. (2012). North-south differences in Spain in iq, educational attainment, per capita income, literacy, life expectancy and employment. *Mankind Quarterly*, 52(3/4):265.

[175] Mackenbach, J. P. (2012). The persistence of health inequalities in modern welfare states: The explanation of a paradox. *Social Science & Medicine*, 75(4):761–769.

[176] Mackenbach, J. P., Bopp, M., Deboosere, P., Kovacs, K., Leinsalu, M., Martikainen, P., Menvielle, G., Regidor, E., and de Gelder, R. (2017). Determinants of the magnitude of socioeconomic inequalities in mortality: A study of 17 european countries. *Health & Place*, 47(C):44–53.

[177] Mackenbach, J. P., Karanikolos, M., and McKee, M. (2013). The unequal health of europeans: successes and failures of policies. *The Lancet*, 381(9872):1125–1134.

[178] Mackenbach, J. P., Kulhánová, I., Artnik, B., Bopp, M., Borrell, C., Clemens, T., Costa, G., Dibben, C., Kalediene, R., and Lundberg, O. (2016). Changes in mortality inequalities over two decades: register based study of european countries. *bmj*, 353:i1732.

[179] Mackenbach, J. P. and Kunst, A. E. (1997). Measuring the magnitude of socio-economic inequalities in health: an overview of available measures illustrated with two examples from europe. *Social science & medicine*, 44(6):757–771.

[180] Mackenbach, J. P., Stirbu, I., Roskam, A.-J. R., Schaap, M. M., Menvielle, G., Leinsalu, M., and Kunst, A. E. (2008). Socioeconomic inequalities in health in 22 european countries. *New England Journal of Medicine*, 358(23):2468–2481.

[181] Manson, J. E., Skerrett, P. J., Greenland, P., and VanItallie, T. B. (2004). The escalating pandemics of obesity and sedentary lifestyle: a call to action for clinicians. *Archives of internal medicine*, 164(3):249–258.

[182] Manton, K. G. (1982). Changing concepts of morbidity and mortality in the elderly population. *The Milbank Memorial Fund Quarterly. Health and Society*, pages 183–244.

[183] Manton, K. G., Stallard, E., and Corder, L. (1997). Education-specific estimates of life expectancy and age-specific disability in the u.s. elderly population 1982 to 1991. *Journal of Aging and Health*, 9(4):419–450.

[184] Manton, K. G. and Tolley, H. D. (1991). Rectangularization of the survival curve: Implications of an ill-posed question. *Journal of Aging and Health*, 3(2):172–193.

[185] Marmot, M. (2002). The influence of income on health: views of an epidemiologist. *Health affairs*, 21(2):31–46.

[186] Marmot, M. (2005). Social determinants of health inequalities. *The lancet*, 365(9464):1099–1104.

[187] Marmot, M. and Bell, R. (2016). Social inequalities in health: a proper concern of epidemiology. *Annals of Epidemiology*, 26(4):238–240.

[188] Marquet, O. and Miralles-Guasch, C. (2015). Neighbourhood vitality and physical activity among the elderly: The role of walkable environments on active ageing in barcelona, spain. *Social Science & Medicine*, 135:24–30.

[189] Martin, L. G., Schoeni, R. F., and Andreski, P. M. (2010). Trends in health of older adults in the united states: Past, present, future. *Demography*, 47(1):S17–S40.

[190] Martín, M. M., López, D. M., and García, G. M. (2014). Estructura ocupacional y cambio urbano en la andalucía oriental del primer tercio del siglo xx. *Revista de Demografía Histórica*, 32(1):73–102.

[191] Mather, M., Jacobsen, L. A., and Pollard, K. M. (2015). *Aging in the United States*, volume 70. Population Reference Bureau.

[192] Mathers, C. D., Stevens, G. A., Boerma, T., White, R. A., and Tobias, M. I. (2015). Causes of international increases in older age life expectancy. *The Lancet*, 385(9967):540–548.

[193] McDade, T. W. and Adair, L. S. (2001). Defining the "urban" in urbanization and health: a factor analysis approach. *Social Science & Medicine*, 53(1):55–70.

[194] McPake, B. and Hanson, K. (2016). Managing the public–private mix to achieve universal health coverage. *The Lancet*, 388(10044):622–630.

[195] Mehta, V. and Bosson, J. K. (2010). Third places and the social life of streets. *Environment and Behavior*, 42(6):779–805.

[196] Michie, S., Abraham, C., Whittington, C., McAteer, J., and Gupta, S. (2009). Effective techniques in healthy eating and physical activity interventions: a meta-regression. *Health Psychology*, 28(6):690.

[197] Mills, M. (2011). *Introducing survival and event history analysis*. Sage Publications.

[198] Ministerio de Empleo y Seguridad Social (2017). Seguridad Social. Accessed: 29/09/2017.

[199] Modin, B. and Fritzell, J. (2009). The long arm of the family: are parental and grandparental earnings related to young men's body mass index and cognitive ability? *International journal of epidemiology*, 38(3):733–744.

[200] Montez, J. K., Hummer, R. A., Hayward, M. D., Woo, H., and Rogers, R. G. (2011). Trends in the educational gradient of u.s. adult mortality from 1986 through 2006 by race, gender, and age group. *Research on Aging*, 33(2):145–171.

[201] Moran, P. A. P. (1948). The Interpretation of Statistical Maps. *Journal of the Royal Statistical Society. Series B (Methodological)*, 10(2):243–251.

[202] Morris, T., Manley, D., and Sabel, C. E. (2018). Residential mobility: Towards progress in mobility health research. *Progress in Human Geography*, 42(1):112–133.

[203] Muszyńska, M. M. and Rau, R. (2012). The old-age healthy dependency ratio in europe. *Journal of population ageing*, 5(3):151–162.

[204] Navarrete, L. (2016). La emigración de los jóvenes españoles en el contexto de la crisis. análisis y datos de un fenómeno difícil de cuantificar. *METAMORFOSIS*, pages 101–105.

[205] Nolasco, A., Melchor, I., Pina, J. A., Pereyra-Zamora, P., Moncho, J., Tamayo, N., García-Senchermes, C., Zurriaga, O., and Martínez-Beneito, M. A. (2009). Preventable avoidable mortality: Evolution of socioeconomic inequalities in urban areas in spain, 1996–2003. *Health & Place*, 15(3):732–741.

[206] Nolasco, A., Moncho, J., Quesada, J. A., Melchor, I., Pereyra-Zamora, P., Tamayo-Fonseca, N., Martínez-Beneito, M. A., Zurriaga, O., Ballesta, M., and Daponte, A. (2015). Trends in socioeconomic inequalities in preventable mortality in urban areas of 33 spanish cities, 1996–2007 (medea project). *International journal for equity in health*, 14(1):33.

[207] Ocaña Riola, R. and Mayoral-Cortés, J. M. (2010). Spatio-temporal trends of mortality in small areas of Southern Spain. *BMC Public Health*, 10(1):26.

[208] Ocaña Riola, R. and Sánchez-Cantalejo, C. (2005). Rurality index for small areas in Spain. *Social Indicators Research*, 73(2):247–266.

[209] OECD (2016). Italy. In *OECD Labour Force Statistics 2015*. OECD Publishing, Paris.

[210] Oeppen, J. and Vaupel, J. W. (2002). Broken limits to life expectancy. *Science*, 296(5570):1029–1031.

[211] Oke, T. (2011). Urban heat island. *The Handbook of Urban Ecology*, pages 120–131.

[212] Oksuzyan, A., Juel, K., Vaupel, J. W., and Christensen, K. (2008). Men: good health and high mortality. sex differences in health and aging. *Aging clinical and experimental research*, 20(2):91.

[213] Oldenburg, R. and Brissett, D. (1982). The third place. *Qualitative sociology*, 5(4):265–284.

[214] Oliva, J. (2010). Rural melting-pots, mobilities and fragilities: Reflections on the Spanish case. *Sociologia Ruralis*, 50(3):277–295.

[215] Olshansky, S. J., Antonucci, T., Berkman, L., Binstock, R. H., Boersch-Supan, A., Cacioppo, J. T., Carnes, B. A., Carstensen, L. L., Fried, L. P., and Goldman, D. P. (2012). Differences in life expectancy due to race and educational differences are widening, and many may not catch up. *Health Affairs*, 31(8):1803–1813.

[216] Olshansky, S. J., Carnes, B. A., and Cassel, C. (1990). In search of methuselah: estimating the upper limits to human longevity. *Science*, 250(4981):634–640.

[217] Omran, A. R. (2005). The epidemiologic transition: a theory of the epidemiology of population change. *The Milbank Quarterly*, 83(4):731–757.

[218] Organization, W. H. et al. (2017). Depression and other common mental disorders: global health estimates. Technical report, World Health Organization.

[219] Owen, N., Sparling, P. B., Healy, G. N., Dunstan, D. W., and Matthews, C. E. (2010). Sedentary behavior: emerging evidence for a new health risk. In *Mayo Clinic Proceedings*, volume 85, pages 1138–1141. Elsevier.

[220] Palloni, A. (2006). Reproducing inequalities: Luck, wallets, and the enduring effects of childhood health. *Demography*, 43(4):587–615.

[221] Palloni, A., Milesi, C., White, R. G., and Turner, A. (2009). Early childhood health, reproduction of economic inequalities and the persistence of health and mortality differentials. *Social science & medicine*, 68(9):1574–1582.

[222] Palmore, E. B., Fillenbaum, G. G., and George, L. K. (1984). Consequences of retirement. *Journal of gerontology*, 39(1):109–116.

[223] Pankratz, V. S., De Andrade, M., and Therneau, T. M. (2005). Random-effects cox proportional hazards model: general variance components methods for time-to-event data. *Genetic Epidemiology: The Official Publication of the International Genetic Epidemiology Society*, 28(2):97–109.

[224] Parker, M. G. and Thorslund, M. (2007). Health Trends in the Elderly Population: Getting Better and Getting Worse. *The Gerontologist*, 47(2):150–158.

[225] Pearce, J., Shortt, N., Rind, E., and Mitchell, R. (2016). Life course, green space and health: Incorporating place into life course epidemiology. *International Journal of Environmental Research*, 13(3):331.

[226] Pérez Díaz, J. (2005). Consecuencias sociales del envejecimiento demográfico. *Papeles de Economía Espãola*, 104.

[227] Permanyer, I., Spijker, J., Blanes, A., and Renteria, E. (2018). Longevity and lifespan variation by educational attainment in spain: 1960–2015. *Demography*, 55(6):2045–2070.

[228] Pickett, K. E. and Wilkinson, R. G. (2015). Income inequality and health: A causal review. *Social Science & Medicine*, 128:316–326.

[229] Pingault, J.-B., O'reilly, P. F., Schoeler, T., Ploubidis, G. B., Rijsdijk, F., and Dudbridge, F. (2018). Using genetic data to strengthen causal inference in observational research. *Nature Reviews Genetics*, 19(9):566.

[230] Pla, A. C. and Cabrerizo, J. A. (2004). Home ownership and social inequality in Spain. *Home Ownership and Social Inequality. Stanford*, pages 233–254.

[231] Portha, B., Chavey, A., and Movassat, J. (2011). Early-life origins of type 2 diabetes: fetal programming of the beta-cell mass. *Experimental diabetes research*, 2011.

[232] Potente, C. and Monden, C. (2018). Disability pathways preceding death in england by socio-economic status. *Population studies*, 72(2):175–190.

[233] Pozuelo, I. (2003). Características del desarrollo urbano reciente en andalucía. *El siglo XX: Geografía de Andalucía. Barcelona, Ariel Geografía*.

[234] Preston, P. (2007). *The Spanish Civil War: reaction, revolution and revenge*. WW Norton & Company.

[235] Rabinowitz, P. M., Slizovskiy, I. B., Lamers, V., Trufan, S. J., Holford, T. R., Dziura, J. D., Peduzzi, P. N., Kane, M. J., Reif, J. S., Weiss, T. R., et al. (2015). Proximity to natural gas wells and reported health status: results of a household survey in Washington County, Pennsylvania. *Environmental Health Perspectives*, 123(1):21.

[236] Radl, J. (2012). Labour Market Exit and Social Stratification in Western Europe: The Effects of Social Class and Gender on the Timing of Retirement. *European Sociological Review*, 29(3):654–668.

[237] Rechel, B., Grundy, E., Robine, J.-M., Cylus, J., Mackenbach, J. P., Knai, C., and McKee, M. (2013). Ageing in the european union. *The Lancet*, 381(9874):1312 – 1322.

[238] Regidor, E., Barrio, G., Bravo, M. J., and de la Fuente, L. (2013). Has health in Spain been declining since the economic crisis? *Journal of Epidemiology and Community Health*, pages jech–2013–202944.

[239] Regidor, E., Vallejo, F., Giráldez-García, C., Ortega, P., Santos, J. M., Astasio, P., and de la Fuente, L. (2015). Low mortality in the poorest areas of spain: adults residing in provinces with lower per capita income have the lowest mortality. *European Journal of Epidemiology*, 30(8):637–648.

[240] Rehnberg, J. and Fritzell, J. (2016). The shape of the association between income and mortality in old age: A longitudinal Swedish national register study. *SSM-Population Health*, 2:750–756.

[241] Rey, G., Jougla, E., Fouillet, A., and Hémon, D. (2009). Ecological association between a deprivation index and mortality in France over the period 1997–2001: variations with spatial scale, degree of urbanicity, age, gender and cause of death. *BMC public health*, 9(1):33.

[242] Ribeiro, W. S., Bauer, A., Andrade, M. C. R., York-Smith, M., Pan, P. M., Pingani, L., Knapp, M., Coutinho, E. S. F., and Evans-Lacko, S. (2017). Income inequality and mental illness-related morbidity and resilience: a systematic review and meta-analysis. *The Lancet Psychiatry*, 4(7):554–562.

[243] Riley, J. C. (2001). *Rising life expectancy: a global history*. Cambridge University Press.

[244] Ritschard, G., Bürgin, R., and Studer, M. (2013). Exploratory mining of life event histories. In *Contemporary issues in exploratory data mining in the behavioral sciences*, pages 243–276. Routledge.

[245] Riva, M., Curtis, S., and Norman, P. (2011). Residential mobility within england and urban–rural inequalities in mortality. *Social science & medicine*, 73(12):1698–1706.

[246] Rodgers, G. B. (1979). Income and inequality as determinants of mortality: an international cross-section analysis. *Population studies*, 33(2):343–351.

[247] Rodríguez, J. S. (2002). Las emigraciones interiores en españa durante los siglos xix y xx: una revisión bibliográfica. *Ager: Revista de estudios sobre despoblación y desarrollo rural= Journal of depopulation and rural development studies*, (2):227–248.

[248] Romanos, E. (2014). Evictions, petitions and escraches: Contentious housing in austerity spain. *Social Movement Studies*, 13(2):296–302.

[249] Roses, J. R. (2004). Why isn't the whole of spain industrialized? new economic geography and early industrialization, 1797–1910. *The Journal of Economic History*, 63(4):995–1022.

[250] Roth, D. L., Fredman, L., and Haley, W. E. (2015). Informal caregiving and its impact on health: A reappraisal from population-based studies. *The Gerontologist*, 55(2):309–319.

[251] Royston, P. and Parmar, M. K. (2002). Flexible parametric proportional-hazards and proportional-odds models for censored survival data, with application to prognostic modelling and estimation of treatment effects. *Statistics in medicine*, 21(15):2175–2197.

[252] Ruiz-Ramos, M. and Viciana-Fernández, F. (2004). Desigualdades en longevidad y calidad de vida entre Andalucía y España. *Gaceta Sanitaria*, 18:260–267.

[253] Rydin, Y., Bleahu, A., Davies, M., Dávila, J. D., Friel, S., De Grandis, G., Groce, N., Hallal, P. C., Hamilton, I., Howden-Chapman, P., Lai, K. M., Lim, C. J., Martins, J., Osrin, D., Ridley, I., Scott, I., Taylor, M., Wilkinson, P., and Wilson, J. (2012). Shaping cities for health: complexity and the planning of urban environments in the 21st century. *Lancet*, 379(9831):2079–108.

[254] Sagardui-Villamor, J., Guallar-Castillón, P., García-Ferruelo, M., Banegas, J. R., and Rodríguez-Artalejo, F. (2005). Trends in disability and disability-free life expectancy among elderly people in Spain: 1986-1999. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, 60(8):1028–1034.

[255] Sánchez, A. R. (2014). The automatic adjustment of pension expenditures in Spain:an evaluation of the 2013 pension reform. Working Papers 1420, Banco de España;Working Papers Homepage.

[256] Sanderson, W. C. and Scherbov, S. (2007). A new perspective on population aging. *Demographic research*, 16(2):27–58.

[257] Sanderson, W. C. and Scherbov, S. (2013). The characteristics approach to the measurement of population aging. *Population and Development Review*, 39(4):673–685.

[258] Santillana, I. (1981). Los determinantes económicos de las migraciones internas en españa, 1960-1973.

[259] Sasson, I. (2016). Trends in life expectancy and lifespan variation by educational attainment: United states, 1990–2010. *Demography*, 53(2):269–293.

[260] Schiaffino, A., Fernandez, E., Borrell, C., Salto, E., Garcia, M., and Borras, J. M. (2003). Gender and educational differences in smoking initiation rates in Spain from 1948 to 1992. *European Journal of Public Health*, 13(1):56–60.

[261] Serrano, J. P., Latorre, J. M., and Gatz, M. (2014). Spain: Promoting the welfare of older adults in the context of population aging. *The Gerontologist*, 54(5):733–740.

[262] Shafey, O., Fernández, E., Thun, M., Schiaffino, A., Dolwick, S., and Cokkinides, V. (2004). Cigarette advertising and female smoking prevalence in spain, 1982–1997: case studies in international tobacco surveillance. *Cancer: Interdisciplinary International Journal of the American Cancer Society*, 100(8):1744–1749.

[263] Siegrist, J. and Wahrendorf, M. (2013). Quality of work, wellbeing, and retirement. *The SAGE handbook of aging, work and society*, pages 314–326.

[264] Sokal, R. R. and Wartenberg, D. E. (1983). A test of spatial autocorrelation analysis using an isolation-by-distance model. *Genetics*, 105(1):219–237.

[265] Solé-Auró, A. and Alcañiz, M. (2015). Are we living longer but less healthy? trends in mortality and morbidity in Catalonia (Spain), 1994–2011. *European journal of ageing*, 12(1):61–70.

[266] Song, J. W. and Chung, K. C. (2010). Observational studies: cohort and case-control studies. *Plastic and reconstructive surgery*, 126(6):2234.

[267] Soobader, M.-J. and LeClere, F. B. (1999). Aggregation and the measurement of income inequality: effects on morbidity. *Social Science & Medicine*, 48(6):733–744.

[268] Speck, B. J. and Looney, S. W. (2001). Effects of a minimal intervention to increase physical activity in women: daily activity records. *Nursing Research*, 50(6):374–378.

[269] Spector, W. D., Katz, S., Murphy, J. B., and Fulton, J. P. (1987). The hierarchical relationship between activities of daily living and instrumental activities of daily living. *Journal of chronic diseases*, 40(6):481–489.

[270] Stein, S. J. and Stein, B. H. (2000). *Silver, trade, and war: Spain and America in the making of early modern Europe*. JHU Press.

[271] Stout, R. W. and Crawford, V. (1988). Active-life expectancy and terminal dependency: Trends in long-term geriatric care over 33 years. *The Lancet*, 331(8580):281–283.

[272] Stuckler, D., Basu, S., Suhrcke, M., Coutts, A., and McKee, M. (2011). Effects of the 2008 recession on health: a first look at european data. *The Lancet*, 378(9786):124–125.

[273] Studer, M. (2013). Weightedcluster library manual: A practical guide to creating typologies of trajectories in the social sciences with r. *LIVES Working Papers*.

[274] Studer, M. and Ritschard, G. (2016). What matters in differences between life trajectories: A comparative review of sequence dissimilarity measures. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 179(2):481–511.

[275] Studer, M., Ritschard, G., Gabadinho, A., and Müller, N. S. (2011). Discrepancy analysis of state sequences. *Sociological Methods & Research*, 40(3):471–510.

[276] Subramanian, S. and Kawachi, I. (2006). Whose health is affected by income inequality? a multilevel interaction analysis of contemporaneous and lagged effects of state income inequality on individual self-rated health in the united states. *Health & place*, 12(2):141–156.

[277] Sullivan, D. F. (1971). A single index of mortality and morbidity. *HSMHA health reports*, 86(4):347.

[278] Szanton, S. L., Seplaki, C. L., Thorpe, R. J., Allen, J. K., and Fried, L. P. (2010). Socioeconomic status is associated with frailty: the women's health and aging studies. *Journal of Epidemiology & Community Health*, 64(01):63–67.

[279] Taylor, M. G. (2010). Capturing transitions and trajectories: the role of socioeconomic status in later life disability. *Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, 65(6):733–743.

[280] Thatcher, A. R., Cheung, S. L. K., Horiuchi, S., and Robine, J.-M. (2010). The compression of deaths above the mode. *Demographic research*, 22:505.

[281] Therneau, T. (2015). Mixed effects cox models. *CRAN repository*.

[282] Therneau, T. M. and Grambsch, P. (2000). Extending the Cox model. *Edited by P. Bickel, P. Diggle, S. Fienberg, K. Krickeberg*, page 51.

[283] Therneau, T. M. and Grambsch, P. M. (2013). *Modeling survival data: extending the Cox model*. Springer Science & Business Media.

[284] Tuljapurkar, S. and Edwards, R. D. (2011). Variance in death and its implications for modeling and forecasting mortality. *Demographic Research*, 24:497.

[285] United Nations, U. (2015). World Urbanization Prospects: The 2014 Revision. Report, Department of Economic and Social Affairs, Population Division.

[286] Vallin, J. and Meslé, F. (2004). Convergences and divergences in mortality: A new approach of health transition. *Demographic Research*, S2:11–44.

[287] Vallin, J. and Meslé, F. (2010). Will life expectancy increase indefinitely by three months every year? *Population & Societies*, (473):1.

[288] Valls, A. D. I., Coll, A. S., and Rivera, E. O. (2014). ¿ migración neohispánica? el impacto de la crisis económica en la emigración española. *EMPIRIA. Revista de Metodología de las Ciencias Sociales*, 29:39–66.

[289] Van de Poel, E., O'Donnell, O., and Van Doorslaer, E. (2012). Is there a health penalty of China's rapid urbanization? *Health Economics*, 21(4):367–385.

[290] Van den Berg, G. J., Doblhammer, G., and Christensen, K. (2009). Exogenous determinants of early-life conditions, and mortality later in life. *Social science & medicine*, 68(9):1591–1598.

[291] Van Oyen, H., Nusselder, W., Jagger, C., Kolip, P., Cambois, E., and Robine, J.-M. (2013). Gender differences in healthy life years within the eu: an exploration of the "health–survival" paradox. *International journal of public health*, 58(1):143–155.

[292] van Raalte, A. A., Kunst, A. E., Deboosere, P., Leinsalu, M., Lundberg, O., Martikainen, P., Strand, B. H., Artnik, B., Wojtyniak, B., and Mackenbach, J. P. (2011). More variation in lifespan in lower educated groups: evidence from 10 european countries. *International Journal of Epidemiology*, 40(6):1703–1714.

[293] van Raalte, A. A., Sasson, I., and Martikainen, P. (2018). The case for monitoring life-span inequality. *Science*, 362(6418):1002–1004.

[294] Van Vliet, D., de Vugt, M. E., Bakker, C., Koopmans, R. T., and Verhey, F. R. (2010). Impact of early onset dementia on caregivers: a review. *International journal of geriatric psychiatry*, 25(11):1091–1100.

[295] Vaupel, J. W., Manton, K. G., and Stallard, E. (1979). The impact of heterogeneity in individual frailty on the dynamics of mortality. *Demography*, 16(3):439–454.

[296] Vaupel, J. W., Zhang, Z., and van Raalte, A. A. (2011). Life expectancy and disparity: an international comparison of life table data. *BMJ open*, pages bmjopen–2011–000128.

[297] Viciana, F., Hernández, J., Canto, V., and Ávila, A. (2003). Longevidad y calidad de vida en andalucía. *Sevilla: Instituto de Estadística de Andalucía*.

[298] Viciana-Fernández, F., Ruiz-Ramos, M., and Pujolar, A. E. (2008). *Andalucía y España. Desigualdades en la Esperanza de Vida*, book section 3. Asociación para la Defensa de la Sanidad Pública de Andalucía.

[299] Vidal-Meliá, C. (2013). An assessment of the 2011 Spanish pension reform using the Swedish system as a benchmark. *Journal of Pension Economics and Finance*, 13(3):297–333.

[300] Vidal-Meliá, C. and Domínguez-Fabián, I. (2006). The Spanish pension system: Issues of introducing ndcs. *Pension Reform*, page 611.

[301] Vlahov, D. and Galea, S. (2002). Urbanization, urbanicity, and health. *Journal of Urban Health*, 79:S1–S12.

[302] Vlahov, D., Galea, S., and Freudenberg, N. (2005). The urban health "advantage". *Journal of Urban Health*, 82(1):1–4.

[303] Waller, L. A. and Gotway, C. A. (2004). *Applied spatial statistics for public health data*, volume 368. John Wiley & Sons.

[304] Walter, S., Beltrán-Sánchez, H., Regidor, E., Gomez-Martin, C., Del-Barrio, J. L., Gil-de Miguel, A., Subramanian, S., and Gil-Prieto, R. (2016). No evidence of morbidity compression in Spain: a time series study based on national hospitalization records. *International journal of public health*, 61(7):729–738.

[305] Warren, J. R., Luo, L., Halpern-Manners, A., Raymo, J. M., and Palloni, A. (2015). Do different methods for modeling age-graded trajectories yield consistent and valid results? *American Journal of Sociology*, 120(6):1809–1856.

[306] Whitehead, M. and Dahlgren, G. (2006). Concepts and principles for tackling social inequities in health: Levelling up part 1. *World Health Organization: Studies on social and economic determinants of population health*, 2.

[307] Wilmoth, J. R. (2000). Demography of longevity: past, present, and future trends. *Experimental gerontology*, 35(9-10):1111–1129.

[308] Wilmoth, J. R. and Horiuchi, S. (1999). Rectangularization revisited: variability of age at death within human populations. *Demography*, 36(4):475–495.

[309] Wise, D. A. (2012). *Social security programs and retirement around the world: Historical trends in mortality and health, employment, and disability insurance participation and reforms*. University of Chicago Press.

[310] Wolf, D. A., Mendes de Leon, C. F., and Glass, T. A. (2007). Trends in Rates of Onset of and Recovery From Disability at Older Ages: 1982–1994. *The Journals of Gerontology: Series B*, 62(1):S3–S10.

[311] Wolfson, M., Rowe, G., Gentleman, J. F., and Tomiak, M. (1993). Career earnings and death: a longitudinal analysis of older Canadian men. *Journal of Gerontology*, 48(4):S167–S179.

[312] Woods, R. (2003). Urban-rural mortality differentials: An unresolved debate. *Population and development review*, pages 29–46.

[313] Woolf, S. H. and Braveman, P. (2011). Where health disparities begin: the role of social and economic determinants—and why current policies may make matters worse. *Health affairs*, 30(10):1852–1859.

[314] Wu, C., Odden, M. C., Fisher, G. G., and Stawski, R. S. (2016). Association of retirement age with mortality: a population-based longitudinal study among older adults in the usa. *Journal of Epidemiology and Community Health*, 70(9):917–923.

[315] Xu, Y. and Wang, F. (2015). Built environment and obesity by urbanicity in the US. *Health & Place*, 34.

[316] Yajnik, C. (2002). The lifecycle effects of nutrition and body size on adult adiposity, diabetes and cardiovascular disease. *Obesity Reviews*, 3(3):217–224.

Appendix A

Chapter 2 - Structural Inequalities and Survival within the Retired Population of Southern Spain

Table A.1 Cox PH Model - Only Individual Public Pension Income

	Dependent variable:	
	Hazard Ratios	
	Male	Female
1000-1999 Eur/month	1.150*** (1.114, 1.186)	1.083 (0.964, 1.218)
650-999 Eur/month	1.148*** (1.113, 1.184)	1.187** (1.056, 1.334)
$<$ 650 Eur/month	1.240*** (1.203, 1.279)	1.043* (0.932, 1.166)
Observations	555,193	276,038
Wald Test (df = 3)	266.890***	45.960***
LR Test (df = 3)	272.132***	44.522***
Note:		*p<0.1; **p<0.05; ***p<0.01

Table A.2 Cox PH Model - Only Household Pension Income

	Dependent variable:	
	Hazard Ratios	
	Male	Female
HH inc. 1000-2000 Euro/month	1.091** (1.039, 1.142)	1.040 (0.953, 1.127)
HH inc. < 1000 Euro/month	1.825*** (1.764, 1.886)	0.736*** (0.638, 0.834)
Observations	113,690	113,684
Wald Test (df = 2)	729.780***	145.120***
LR Test (df = 2)	663.719***	153.127***
Note:		*p<0.1; **p<0.05; ***p<0.01

Appendix B

Chapter 3 - The effects of socioeconomic position on old age mortality within shared disability pathways

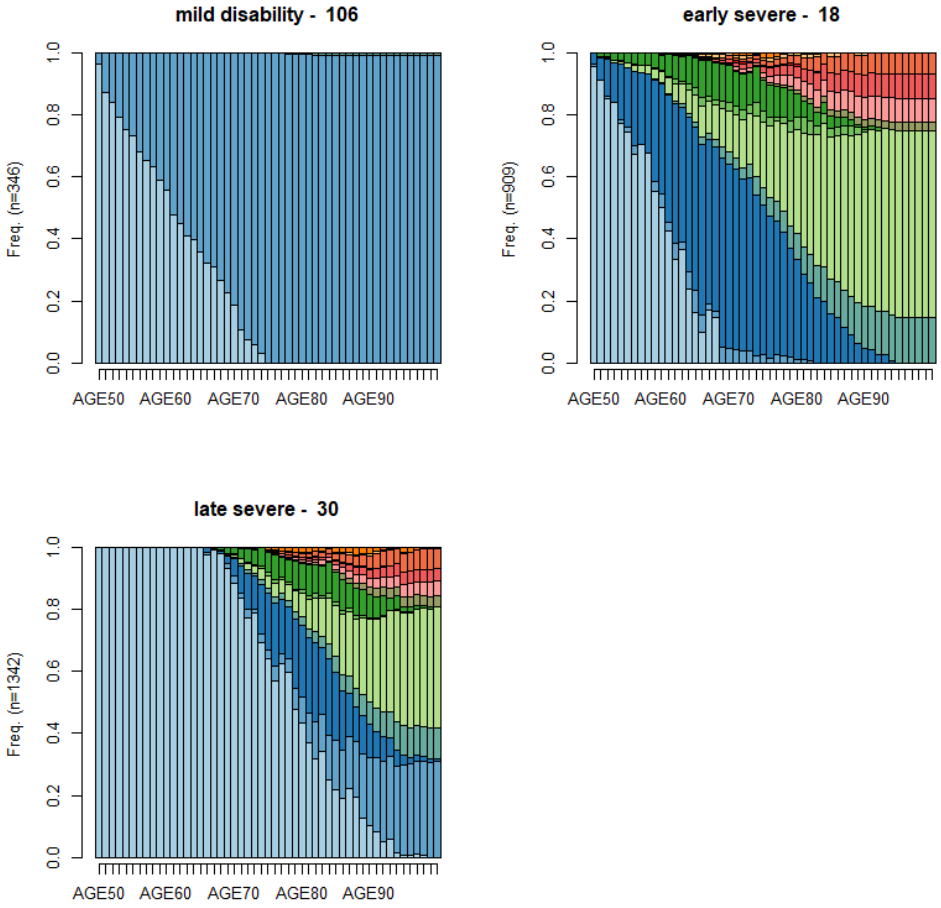


Fig. B.1 Relative Distribution of State Sequences by Pathway and Age (Males)

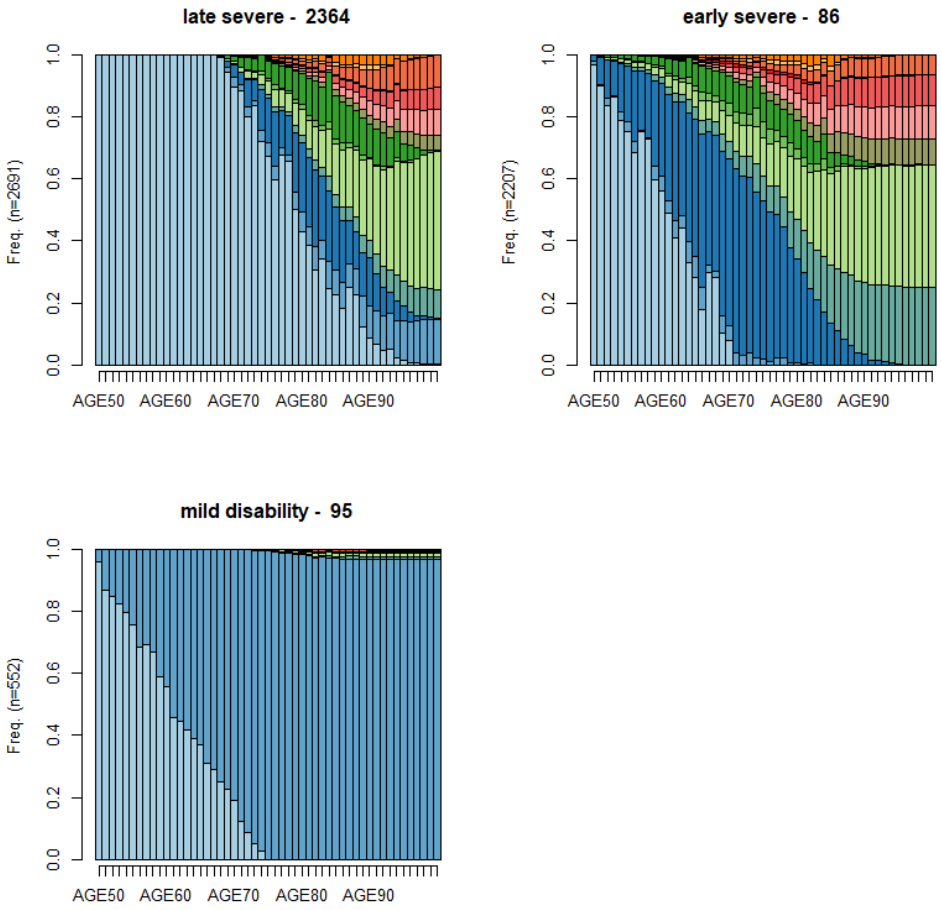


Fig. B.2 Relative Distribution of State Sequences by Pathway and Age (Females)

Table B.1 Gompertz PH Regression Models - Females Mild Disability Trajectory

	Dependent Variable: Relative risk of dying		
	(1)	(2)	(3)
No daily activity	1.506*	1.444*	1.420
	(1.067, 2.124)	(1.018, 2.047)	(0.995, 2.025)
Reference: Daily active			
Suffers from multiple diseases	1.325	1.336	1.327
	(0.8500, 2.066)	(0.853, 2.092)	(0.845, 2.082)
Reference: No multi-morbidity			
Disability Onset 0-2 years	1.316	1.309	1.370
	(0.805, 2.151)	(0.797, 2.147)	(0.826, 2.270)
Disability Onset 3-6 years	1.810+	1.909*	1.998*
	(1.077, 3.042)	(1.124, 3.243)	(1.161, 3.438)
Disability Onset 7-12 years	0.816	0.822	0.855
	(0.509, 1.308)	(0.510, 1.327)	(0.525, 1.391)
Reference: 12+ *years*			
Income per CU $<$ 500 Euro		1.469	1.463
		(0.942, 2.291)	(0.929, 2.306)
Income per CU 500-750 Euro		1.450	1.335
		(0.947, 2.219)	(0.827, 2.154)
Reference: $>$ *2000 Euro*			
Incomplete Educ.		0.726	0.785
		(0.407, 1.291)	(0.425, 1.354)
Primary Educ.		0.676	0.681
		(0.363, 1.261)	(0.365, 1.270)
Reference: Secondary/Higher Educ.			
Widowed			0.804
			(0.502, 1.288)
Div./Single			0.949
			(0.495, 1.820)
Reference: Married			
No Close Relatives			1.290
			(0.471, 3.535)
Close kin in same HH			1.047
			(0.516, 2.121)
Close kin in same municipality			0.841
			(0.488, 1.451)
Reference: Kin in other geogr. area			
$>$ 2 people in HH			1.068
			(0.569, 2.002)
Reference: 1-2 people in HH			
Observations	515	515	515
Log Likelihood	−562.808	−560.463	−559.073
LR Test (df= total/additional)	13.632 (df = 7/7)	4.690 (df = 11/4) **	2.779 (df = 17/6)
AIC	1139.616	1142.926	1152.146
Note:			+p<0.1; *p<0.05; **p<0.01; ***p<0.001

Table B.2 Gompertz PH Regression Models - Males with Mild Disability Trajectory

	Dependent variable: Relative risk of dying		
	(1)	(2)	(3)
No daily activity	1.655**	1.588**	1.630**
	(1.169, 2.343)	(1.117, 2.260)	(1.143, 2.324)
Reference: Daily Active			
Suffers from multiple diseases	1.221	1.157	1.220
	(0.800, 1.866)	(0.756, 1.768)	(0.792, 1.878)
Reference: No Co-morbidity			
Disability Onset 0-2 years	1.769*	1.870*	2.039*
	(1.073, 2.917)	(1.126, 3.106)	(1.208, 3.442)
Disability Onset 3-5 years	1.485	1.430	1.504
	(0.819, 2.690)	(0.788, 2.595)	(0.820, 2.759)
Disability Onset 6-10 years	1.227	1.184	1.208
	(0.742, 2.028)	(0.715, 1.961)	(0.725, 2.015)
Reference: >10 years			
Income per CU < 500 Euro		1.361	1.344
		(0.816, 2.271)	(0.783, 2.306)
Income per CU 500-750 Euro		0.832	0.845
		(0.548, 1.262)	(0.549, 1.301)
Reference: > 750 Euro			
Incomplete Educ.		0.968	0.960
		(0.613, 1.150)	(0.605, 1.525)
Primary Educ.		0.657	0.652
		(0.396, 1.090)	(0.392, 1.084)
Reference: Secondary/Higher Educ.			
Widowed			0.714
			(0.374, 1.363)
Div./Single			1.690
			(0.926, 3.083)
Reference: Married			
No Close Relatives			0.984
			(0.364, 2.659)
Close kin in same HH			1.438
			(0.682, 3.033)
Close kin in same municipality			1.215
			(0.715, 2.064)
Reference: Kin in other geogr. area			
> 2 people in HH			1.111
			(0.560, 2.206)
Reference: 1-2 people in HH			
Observations	305	305	305
Log Likelihood	−494.815	−487.624	−488.041
LR Test (df= total/additional)	13.849+ (df = 7)	8.021+ (df = 11)	5.526 (df = 17)
AIC	1003.63	1003.61	1010.08

Note: +p<0.1; *p<0.05; **p<0.01; ***p<0.001

Appendix C

Chapter 4 - Urban environment and mortality disparities in Andalusia

Table C.1 Factor loadings and uniqueness parameter from the maximum likelihood factor analysis

	Factor Loading	*Uniqueness Parameter*
Population Density	0.97	0.05
% of Artificial Surface Area	0.82	0.33
% of Health Service Area	0.90	0.19
Road Density	0.65	0.58
Note:	Estimated factor explains 71% of the variance.	

Table C.2 Morans I Statistics for univariate relationships of indicators with geographic dimension

	Morans I Statistic	*p-value*
Smoothed SMR	0.380	< 0.001
"Urbanicity" Index	0.784	< 0.001
Perceived Noise	0.570	< 0.001
Perceived Neighborhood Pollution	0.444	< 0.001
% Single Household	0.568	< 0.001
% Unemployed	0.500	< 0.001

Table C.3 Random effects (RE) statistics and model comparison to model without random effects

	Model 1	Model 2	Model 3	Model 4
Standard deviation RE	0.1559	0.1251	0.1202	0.1125
Variance RE	0.0243	0.0156	0.0144	0.0127
AIC	1085414	1077044	1077001	1076938
Log-Likelihood Ratio Test	127.11***	56.36***	48.36***	37.55***
Chi-Square (df)	(1)	(1)	(1)	(1)
Note:		*p<0.01; **p<0.005; ***p<0.001		